Columbia University

Contributions to Education

Teachers College Series

SPELLING ABILITY

ITS MEASUREMENT AND DISTRIBUTION

B. R. BUCKINGHAM, PH.D.

TEACHERS COLLEGE, COLUMBIA UNIVERSITY
CONTRIBUTIONS TO EDUCATION, No. 59

PUBLISHED BY
Teachers College, Columbia University
NEW YORK CITY
1913

CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

SPELLING ABILITY—ITS MEASUREMENT AND DISTRIBUTION

§ 1. *Introduction*

The purpose of this dissertation is to derive a scale for the measurement of spelling ability and to show some of its uses and applications. Such a purpose relates itself closely to a general movement, which is now well under way, and which aims to place in our hands the means of stating with something approaching the precision of objective measurement the amounts of each school ability possessed by an individual or a group. We received not long ago a scale for Handwriting (Thorndike, E. L., 1910) and still more recently a scale for English Composition (Hillegas, Milo B., 1912). The former consists in the use of selected specimens of handwriting each of which has been evaluated; the latter consists of a similar series of English compositions. It will be noticed that some of the conditions of objective measurement are met. We measure given specimens of handwriting by comparing them with actual samples of handwriting of known value. We determine the quality of English composition by a like comparison with samples of actual English writing of known value.

It seems clear, therefore, that if we are to measure ability in spelling at all it will be by reference to an evaluated standard or sample of spelling. If we can arrange a series of words on a linear projection in such a way that the words from the low end to the high end are placed at equal intervals determined by the difficulty of each word, and if we can determine a zero-point such that failure to spell the word fixed at that point under the required conditions indicates *absence* of spelling ability, then we shall have constructed a scale by which we may measure the spelling ability of an individual, or by which we may through suitable tests determine the difficulty of any word in the language. Since the spelling of individuals may

thus be rated, the spelling of classes, of schools, and of school systems may likewise be rated.

It may be said that we have always rated pupils in spelling; and that schools and school systems have likewise been rated. Such is indeed the case. But there has always been a lack of precision in these ratings due to the inequality of the units employed. Dr. Rice (Rice, '97), for example, in testing the pupils in 4th to 8th grades in twenty-one school systems used a list of words containing among others: *disappoint, necessary, changeable, better, because, picture.* The method of rating pupils was the usual one of deducting from 100 per cent the same per cent for each word. That is, all words were taken as equal measures of spelling ability. A moment's attention to the six words mentioned will lead us to suspect that this is not a true assumption; and an actual test of a group of 5th-year children with these words shows that our suspicion is correct. In such a test mistakes were made as follows:

disappoint, 37
necessary, 42
changeable, 42
better, 3
because, 1
picture, 0 (Thorndike, '04, p. 8)

To give these words equal weight in any test is to make inaccurate most of the conclusions based upon it. A pupil who spells all or nearly all of the list is a much better speller than the figures show; for he has probably spelled not only all the easy words but also many of the hard ones. On the other hand, a pupil who misses most of the words is a much poorer speller than his rating indicates because he has probably failed to spell all the hard words as well as most of the easy ones.

Nor is this list of Dr. Rice's at all unusual. Cornman used the same list (Cornman, '02). Both used a composition test where pupils were rated according to the per cent of their correctly spelled words among the total number of words in a written exercise. Cornman also used a test in which school

children were required to write "as many words as they could" in 15 minutes. Of course in the composition test and in the 15-minute test no two children wrote the same words. Moreover, the words written by each child must have varied widely in difficulty. The result for the 15-minute test, according to Cornman's best table, is as follows:

School Year	Median Percentage	Average Variation
8th	97.9	.60
7th	96.2	.50
6th	95.2	.33
5th*a*	94.3	.36
5th*b*	94.3	.10
4th*a*	94.7	.66
4th*b*	93.7	.96
3d*a*	93.5	.23
3d*b*	93.0	1.43

One conclusion from this is that "pupils of the elementary school increase regularly from grade to grade in accuracy of spelling." This might almost be taken for granted. But in answer to the question, "How much does one grade surpass another?" the figures afford no information. Obviously from all we know of the elementary school, the difference between eighth-grade ability and low third-grade ability in spelling is far greater than the figures 97.9 and 93 indicate.

Similarly the Composition Tests of Rice and Cornman are misleading when used to indicate spelling ability. According to the series of Composition Tests of the latter, 8th-year children on the average spelled 99.5 per cent of their words correctly, and children of the first half of the 3d year spelled 93.2 per cent correctly. The author draws conclusions from his figures as to the progress of each grade for the school year, as to the progress of the school and as to the effect of the suspension of instruction in spelling. Since in the series of eight tests the children wrote various kinds of lessons—Geography, History, Science, Language, Composition—each with its own peculiar words, and since each pupil used his own individual vocabulary, we cannot escape the conviction that while these figures may be suggestive of progress, or of the effect of change in method, or of grade differences, they are nothing more than suggestive.

They leave unanswered the questions,—How much progress? How large an effect? How great a difference? As we grow more and more accustomed to quantitative thinking in our educational work, we feel that these are precisely the questions that we ought in some way to be able to answer.

These studies of spelling made by Cornman and Rice remain the most important statistical treatment of the subject. That they have not great value it would be presumptuous even to imply. Their results are in a general sense true. To a certain extent their lists, even though made up of words of various and undetermined difficulty, may be used, especially for comparative purposes, as a total measuring device. They do, however, undoubtedly suffer through lack of precision, while their statements of *amounts* of difference are in general misleading.

The same thing may be said of later investigations. For example, Wallin's tables and his conclusions from them as to the transfer of spelling efficiency and its relation to age, grade, and sex are subject to the same limitations (Wallin, '11). Pearson's "Experimental Studies in the Teaching of Spelling" (Pearson, '12), however, shows a recognition of the difficulty, although it offers no remedy. In his treatment of the relative values of the "together-method" and the "separate-method" of teaching homonyms this author says: "Owing to the inequality of the units of measurements, it is impossible to determine accurately from Table IV whether the together-method is superior to the separate-method. One cannot decide, for example, positively whether an improvement from 3.78 errors (median of a class) to 2.86 errors is greater or less than an improvement from 5.6 errors to 3.3 errors." If, however, the words used could have been evaluated through an independent test by reference to a scientifically constructed scale, the "inequality of the units of measurement" would have disappeared. The further treatment of the foreshortening of the opportunity for improvement due to high initial performance is quite another matter.

It will be clearly seen from the foregoing that in practically all work which has attempted to present the spelling situation statistically it is assumed as fundamental that one error equals

another and that to spell one word is the same as to spell another word.

It will therefore be profitable to seek in this field as others have sought in other fields to devise an instrument which will more accurately measure that of which we are so often called upon to give a quantitative statement.

§ 2. *Limitations*

The study here attempted is confined to the elementary school entirely. It covers the grades from the third to the eighth, both inclusive. The schools tested are all located in or near New York City. The cosmopolitan character of the population of the metropolitan area makes it extremely unlikely that results of a materially different character would have been obtained by testing schools in various sections of the country.

It is believed that these schools are fairly typical within the limits of the area chosen. School I is a private school of high class whose pupils are mostly American born and from good homes. All the other schools are public schools. School II is in a German section of rather low class. School III is in a better neighborhood, foreigners predominating. School IV is in an Italian section. It has long had the benefit of high-class supervision and organization. School V is again predominantly American. It is located outside of the city system. School VI is in a good residential section of the city. School VII is a large school, most of whose pupils are of foreign parentage. Territorially, two schools are in Manhattan, one in the Bronx, one in Brooklyn, and two in Queens, while one is outside of the city entirely.

In all 8,791 pupils were tested. It is thought that this is a sufficient number for practical purposes. In fact it was found that the returns from each additional school after the first three or four made almost no change in the results. It is probable that greatly increasing the number of pupils tested would have afforded little compensation for the additional labor. It has seemed wiser to limit the number to a moderate one and to spend considerable effort in making the statistical analysis as complete as possible.

§ 3. *The Original List*

The preliminary testing was made with a list of 270 words. It will be called "The Original List." It was itself selected from a much larger list of graded words used by the author of this dissertation in his own school, the same having been secured by taking from five of the popular Spelling Books now in use a vocabulary of 5,000 words agreed upon by two or more of the books. The principles of selection for these 270 words were: (1) that all of them should be sufficiently common to be in the speaking vocabulary of third-grade children; and (2) that the spelling difficulty of many of them should be great enough to test the ability of eighth-grade children. As a matter of fact, the selection did not consist of 270 words at first. The list grew to that number only after the chosen words were put into sentences. The necessary helping words then swelled the total to the number given.

The sentences were dictated during the fall term of 1910 to schools I and II. They were given to grades 3 to 8 in School II, and to grades 4 to 7 in School I. Their dictation consumed several periods for every class. The following are the sentences:

There were forty birds on the bridge. Do not go until I come. On *Wednesday an umbrella was found. Whose pencil is this? My uncle gave me a banana.* The *butcher* gave the *hungry dog* a *piece of meat.* My *answer* is *ninety. For* a *nickel* I *bought* an *orange,* a *peach, and* a *pear.* A *dollar* is not *too much money* for *so beautiful* a *picture. Learn to* do *right because it* is right. The *chicken ran across* the *road.* The *janitor sweeps every Tuesday afternoon.* It is *wrong* to *steal even* a *penny.* It *would be easy* to *watch* for *your cousin from* the *parlor window.* It is the *hour* for *recess. Smoke* was *coming out* of *their chimney. One summer evening* my *neighbor came into* my *kitchen.* I *did* not *know he* was coming *that night.* To *whom does* this *pair* of *scissors belong? I am almost sure they* belong to the *tailor.* The *doctor thought* he *ought* to go *at once. His bicycle* was *against* the *fence. But* a *carriage* was *stopping in front* of his *office.* His *friend* was *already beginning* to *speak* to *him.* He *said* the *soldier should have medicine* this *minute. Pshaw,* there was *neither* a *monkey nor* an *elephant* at the *circus. Get some*

coffee, sugar, and *soap* at the *grocery store.* The soldier *dropped* his *sword* and *pistol.* Jack *had* a *whistle* and *nineteen nails* in his *pocket.* The *pretty fairy* had a *saucy tongue.* One *day* in *February people saw* a *sleigh pass through* the *avenue. Shoes are made* of *leather* and a *little iron.* A *week* from *to-day* there *will* be a *dance. Cut up* a *tomato* and an *onion together.* In my *garden* I *shall raise cabbage instead* of *beets.* The *saucer* was *round like* a *circle. Make no noise;* do not *whisper or laugh. Nobody* should be *without* a *handkerchief.* A *straight line* has *length only. We* shall *believe* the *truth.* We have *another piano* at *our school.* Is it *true* that there was *grease* on the *towel?* This *animal* has a *large mouth.* It is not *often cold enough* for the *ocean* to *freeze. Guess what* made me *sneeze. Choose which* one of the *pigeons you* like. *Touch* the *button with* your *thumb.* The *American Indian* had *corn* and *tobacco.* I have *written* the *whole alphabet.* I *wear* a *number thirteen collar. If* the *men quarrel, telephone* me or *send* a *telegram.* Our *arithmetic lesson* is in *addition.* We *also subtract.* A *handful* of corn was *all* I had for *supper.* What is the *title* of the *story?* Did you *hear* the *thunder last* night? I am *tying* up my shoe. A *basin* of *water sat* on the *table.* That *sentence* has *twelve words* in it.

Those who dictated the sentences were directed to read them in whole or in part as many times as seemed necessary to secure their complete comprehension. Pupils were therefore not required to retain in mind a long series of words.

In rating the papers only the words printed in italics were considered. If a word occurred twice it was regarded only the first time it appeared. Omitted and illegible words were classed as wrong. All the papers here as well as elsewhere throughout this study were rated by the same person. They were rated from two points of view: (1) as to the number of times each word was correctly spelled, and (2) as to the per cent of the entire number of words each pupil spelled correctly. The former point of view is the only one to which attention is now directed.

Table I is a sampling from the entire 270 words as given to schools I and II. At School I the grammar-school course is completed in seven years. It therefore has no 8th grade. As stated above, the test was not given to the 3d grade in this school.

TABLE I

FIGURES INDICATE PER CENT CORRECT

Table reads: "across" was spelled correctly in the 3d grade of School II by 17% of the pupils; in the 4th grade of School I by 60% of the pupils, and of School II by 40% of the pupils, etc.

GRADE.....	3d	4th		5th		6th		7th		8th
SCHOOL.....	II	I	II	I	II	I	II	I	II	II
across.....	17	60	40	76	58	90	79	98	87	93
addition.....	2	38	26	60	28	76	45	94	76	83
almost.....	16	62	41	73	65	88	75	80	81	87
alphabet.....	25	13	1	63	12	40	46	82	43	68
arithmetic....	27	89	53	100	72	96	92	100	97	98
bridge.....	29	59	42	87	52	98	85	100	94	97
button.....	14	50	35	70	49	77	63	84	62	83
choose.....	6	25	10	37	31	62	37	67	55	65
day.....	97	100	98	96	100	100	99	100	100	100
guess.....	6	29	17	67	30	77	50	82	66	85
handful.....	36	47	33	46	19	76	33	75	63	57
pshaw.....	1	4	6	29	6	46	5	31	31	18
tomato.....	34	83	49	67	43	74	48	79	32	38
too.....	0	10	3	17	4	26	7	63	22	27
whose.....	17	49	15	40	29	47	10	57	59	66

§ 4. *The Selected List*

On the basis of the results for the Original List, a group of 100 words was chosen. It is here called the "Selected List." In Table I are shown 15 words from the Original List. The word "across" is typical of the words taken for the Selected List. Since 17 per cent of the 3d-grade children spelled it correctly, it was not so difficult in that grade as to offer no test of ability. It showed a steady increase throughout the following grades but did not reach so high a figure in the highest grades as to prevent its being a test of ability there. "Almost" and "button" were chosen for the same reason. "Addition" was not taken because it was too hard for 3d-graders. Only 2 per cent wrote it correctly. So small a number as two in a hundred might get it right by chance. Practically, therefore, the word is a zero word for the 3d grade; and such a word does not test ability. There may be—and in a given grade there certainly would be—wide differences in spelling ability, but such a word

will not show them. "Alphabet" was rejected because though high in the 3d grade it was very low in the 4th, suggesting that in School II it was a word that the children had recently studied. "Arithmetic" was not taken because from the 6th grade on it offered practically no difficulty. As in the case of a word rated at zero or nearly zero, so in the case of a word rated at 100 or nearly 100, there is no test. Good spellers and poor spellers so far as the particular word is concerned behave exactly alike. "Bridge" was not taken for the same reason. "Choose" was too hard in the 3d grade. "Day" was too easy everywhere. In fact "day" is a type of word such that we may almost be warranted in saying that one who cannot spell it has no spelling ability. "Guess" was taken because although it is very seldom spelled correctly in the 3d grade, its form is so peculiar that the few who did write it correctly probably knew how to spell it, i.e., did not get it right by chance. "Handful" is a type of word taken because although it shows no regular increase from grade to grade it offers a real test for every grade. The later results in other schools, however, showed that its irregularity is not accidental in schools I and II but is a peculiarity of the word itself. "Pshaw" is a familiar word to the ear, but not to the eye. Very few get it right in any grade. It was rejected. "Tomato" is curious. On the whole neither school does any better with it in the highest than in the lowest grades. It was not taken. This word and the word "handful" strongly suggest the need of a greater number of pupils to test. The word "too" is a word which is misspelled with astonishing frequency. The difficulty is not so much one of spelling as of confusion with the other two words which have the same pronunciation. It was not used in the Selected List but was later included in a small supplementary list just to "try it out." "Whose" was taken although it shows a dip in the 5th and 6th grades. Pupils in these grades have learned the use of the apostrophe and their "little knowledge" proves a "dangerous thing" which the pupils of the earlier grades avoid.

These words—each more or less typical in its way—show how from the Original List of 270 a better Selected List of 100 was chosen. Again the words were put into sentences, as follows:

Whose answer is *ninety?* If the *janitor sweeps,* he will *raise* a dust. You *ought* not to *steal even* a *penny.* Wait *until* the *hour* for *recess* to *touch* the *button. Smoke* was *coming* out of *their chimney. Every afternoon* the *butcher* gave the *hungry* dog a *piece* of meat. One *evening* a *carriage* was *stopping* in *front* of my *kitchen.* I *wear* a *number thirteen collar. Guess* what made me *sneeze. Send* me a *pair* of *leather* shoes. I do not *know,* but I am *almost sure* they are mine. My *uncle bought* my *cousin* a *pretty watch* for *forty dollars.* The *soldier dropped* his *sword.* Jack had a *whistle* and *also twelve nails.* The *ocean does* not *often freeze.* You *should speak* to *people whom* you meet. It takes *only* a *minute* to *pass through* the gate and *across* the *road.* Did you ever *hear* a *fairy laugh?* The *American Indian* had a *saucer without* a cup. *Neither* a *pear nor* a peach was at the *grocery* store *to-day.* Cut up a *whole onion* with a *handful* of beans. My *piano lesson* was *easy.* The *animal* ran *into* the road and *straight against* a tree. Give me *another sentence which* has the word *"title"* in it. I *believe true friends* like to be *together instead* of apart.

These sentences were dictated at schools III, IV, and V in the spring term of 1911. They were later (fall of 1912) dictated at schools VI and VII.

The following instructions were given to the examiners:

Please read these instructions through before beginning to dictate the sentences.

I. See that *each sheet* is headed with (a) the pupil's name, (b) the date, (c) the grade, (d) the name of the school.

II. Give all the sentences during one session, i.e., either in the morning or the afternoon of the same day.

III. In classes below the fifth year dictate in two periods, separated by at least half an hour, or by a recess period.

IV. Each sentence may be dictated, either in whole or in part, as many times as may seem necessary to secure its complete understanding. This exercise is purely a test in spelling. It is not intended that pupils should be subjected to the added difficulty of an effort to recall the words dictated.

V. Offer no explanation of words or sentences. If the meaning is not clear, repeat the sentence as a whole or in part.

VI. Do not ask the children to underline words nor otherwise call their attention to the significant words of the sentences.

VII. After the children have written the sentences, read them again and allow pupils to insert words or make other corrections.

VIII. Collect the papers.

Subsequently at the same schools (III, IV, and V) was given a supplementary list of 18 words, again selected from the Original List (270 words). With the same directions to the examiners, these words were put into sentences as follows:

Telephone me on *Tuesday* if the *tobacco* comes. The *tailor* sent a *saucy telegram. Already* the *circus* was *beginning. Pigeons* seem *too beautiful* to *quarrel.* I am *trying* to *choose* a *towel.* The *chicken* was fried in *grease.*

Each of these 118 words was scored in each grade and for each school separately. Table II illustrates for a few of the words the manner in which this was done. The figures indicate per cent correct. 3*a* means third grade, first half; 3*b* third grade, second half, etc. III, IV, and V refer to different schools.

It will be seen at once that there is no steady progression for each word as we pass from the lowest to the highest grades. In fact for this and for other reasons it was found best to deal with grades by years rather than by half-years. It also seemed advisable to choose a few of these words and to make them the basis of study.

§ 5. *The Preferred Lists*

From the data now in hand it was possible to select a few words which showed reasonably regular increase from grade to grade in the per cent of times they were spelled correctly. Two lists were made up, each containing twenty-five words. The first list is superior to the second in the testing power of the words in all grades and in the permanence of their relative difficulty throughout the grades. That is, to a somewhat greater extent than in the case of the second list, the words of the first list are found to be easy enough for low grades and hard enough for high grades. Also, a word occupying a certain serial position (say the 4th in point of difficulty for the third grade) tends more strongly in the first than in the second list to occupy the same position in all other grades. That both lists, however, are reasonably satisfactory in these particulars will be abundantly shown.

TABLE II

FIGURES INDICATE PER CENT CORRECT

GRADE	3a	3b	4a	4b	5a	5b	6a	6b	7a	7b	8a	8b
against												
III	20	60	54	43	60	62	60	91	73	93	97	95
IV	5	19	26	29	70	61	92	78	95	94	100	97
V	0	0	10	15	30	52	54	74	63	81	88	89
believe												
III	8	77	37	54	53	44	35	57	73	93	73	85
IV	0	3	17	36	32	22	50	52	62	78	83	78
V	0	4	24	14	28	55	76	50	77	71	84	72
cousin												
III	12	16	44	72	78	83	91	95	98	97	100	100
IV	10	39	38	58	72	82	97	93	97	96	100	100
V	11	28	45	38	55	68	84	74	68	90	86	91
know												
III	22	40	63	75	73	70	88	82	93	97	97	100
IV	37	61	68	71	78	81	85	93	97	98	98	100
V	33	38	62	60	60	84	80	79	79	88	100	96
ninety												
III	56	61	65	46	60	65	70	86	71	80	97	80
IV	15	25	54	19	59	49	61	60	87	82	83	86
V	0	11	28	14	36	52	64	61	62	71	79	74
pigeons												
III	0	2	13	31	24	61	66	49	83	64	96	86
IV	8	6	74	50	63	22	57	67	81	74	84	95
V	0	45	9	32	61	50	82	64	50	89	80	84
saucer												
III	7	4	17	52	45	40	33	73	78	83	80	75
IV	10	10	16	36	49	51	73	60	83	72	78	78
V	0	16	31	19	21	52	66	45	66	79	81	78
too												
III	0	0	15	58	64	46	14	60	24	14	65	32
IV	3	45	34	27	54	40	10	11	33	33	53	45
V	0	45	0	85	18	26	36	41	18	26	51	39

The first list will be called the "First Preferred List." It contains the following words:

1. even
2. lesson
3. only
4. smoke
5. front
6. sure
7. pear
8. bought
9. another
10. forty
11. pretty
12. wear
13. button
14. minute
15. cousin
16. nails
17. janitor
18. saucer
19. stopping
20. sword
21. freeze
22. touch
23. whistle
24. carriage
25. nor

The second list, called the "Second Preferred List," is as follows:

1. already
2. beginning
3. chicken
4. choose
5. circus
6. grease
7. pigeons
8. quarrel
9. saucy
10. tailor
11. telegram
12. telephone
13. tobacco
14. too
15. towel
16. Tuesday
17. tying
18. whole
19. against
20. answer
21. butcher
22. guess
23. instead
24. raise
25. beautiful

Table III gives for each word of the First Preferred List and for each grade the number of times the word was written, the number of times it was spelled correctly, and the per cent correct. Schools I, II, III, IV, and V are included. (Omitted words are considered as "written" and as wrong.) Table IV gives the same facts for the Second Preferred List.

It will be seen from tables III and IV that for any given word the per cent correct in one grade is higher than it is in any lower grade. This is, of course, to be expected. But it is not sufficient. In order that this list should be of greatest value it should be so constituted that these increases in 'per-cents-correct' so keep pace with the increase from grade to grade of general spelling ability that a word tends in all grades to maintain the same difficulty relative to all other words in the list to which it belongs. A word which is 20th in point of difficulty for the 3d grade ought to deviate as little as possible from the same rank in the other grades. The experience gained in making this investigation leads us to think that most words do not meet this condition even approximately. The span between the 3d and the 8th grades is very wide. Accordingly a very large class of words is impossible for the earlier, yet easy for the later, grades. Still others are really difficult in the lower grades but of almost no difficulty in the upper grades. From our own Original List "coffee" and "people" are hard for 3d- and 4th-graders, but are almost always spelled correctly above the 6th grade. A third group of words breaks down in the middle. They appear to be easy in the lowest and highest

TABLE III

First Preferred List

Grade	3d Year				4th Year				5th Year				6th Year				7th Year				8th Year			
	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank
even	426	253	59	3	485	384	79	1	537	476	89	1.5	459	428	93	3	395	368	93	7	267	259	97	4
lesson	421	154	37	12	478	345	72	3.5	531	440	83	4	456	416	91	4.5	397	374	94	4.5	267	256	96	7
only	423	275	65	1	485	365	75	2	535	476	89	1.5	459	434	95	1	397	385	97	1	267	264	99	1.5
smoke	426	197	46	7	485	336	69	5	537	459	85	3	459	432	94	2	395	380	96	2	267	263	99	1.5
front	424	215	51	4	485	348	72	3.5	538	431	80	5	458	414	90	6.5	388	365	94	4.5	267	258	97	4
sure	425	200	47	6	487	267	55	11.5	538	372	69	10.5	458	365	78	14.5	388	345	89	10.5	267	252	94	12
pear	421	129	31	16	486	206	42	20.5	536	309	58	19.5	453	327	72	19.5	397	321	81	19	267	250	94	12
bought	425	169	40	11	489	320	65	7	536	422	79	6	453	414	91	4.5	396	374	94	4.5	267	260	97	4
another	421	151	36	13	483	206	43	19	535	419	78	7	459	393	86	10	397	372	94	4.5	267	255	96	7
forty	425	209	49	5	489	301	62	8	536	346	65	13.5	453	324	72	19.5	396	328	83	17	267	231	87	20.5
pretty	425	190	45	8.5	464	313	67	6	539	411	76	8	460	414	90	6.5	397	358	90	9	267	251	94	12
wear	424	149	35	14	483	236	49	15	531	322	61	16.5	456	336	74	17	396	331	84	15.5	267	249	93	16
button	426	135	32	15	461	238	52	13.5	537	328	61	16.5	459	336	73	18	394	293	74	25	267	232	87	20.5
minute	423	109	26	19	484	185	38	24	538	331	62	15	458	351	77	16	388	334	86	13	267	245	92	18
cousin	425	81	19	21.5	488	227	47	16	537	370	69	10.5	459	409	89	8	395	353	89	10.5	267	254	95	9
nails	425	184	43	10	464	268	58	10	539	385	71	9	460	398	87	9	398	368	92	8	267	256	96	7
janitor	426	82	19	21.5	485	202	42	20.5	537	309	58	19.5	459	373	81	12.5	395	319	81	19	267	239	90	19
saucer	422	48	11	25	485	141	29	25	535	225	42	25	459	268	58	25	397	312	79	21	267	215	81	25
stopping	424	113	27	18	485	191	39	23	538	294	55	23	458	326	71	21	388	296	76	23	267	225	84	24
sword	425	56	13	23.5	464	213	46	17.5	539	308	57	21	460	361	78	14.5	398	343	86	13	267	247	93	16
freeze	425	185	29	17	464	262	46	17.5	537	363	68	12	459	379	83	11	395	340	86	13	267	251	94	12
touch	426	191	45	8.5	461	242	52	13.5	537	324	60	18	459	370	81	12.5	394	331	84	15.5	267	247	93	16
whistle	425	94	22	20	464	254	55	11.5	539	304	56	22	460	294	64	24	398	299	75	24	267	226	85	22.5
carriage	424	55	13	23.5	485	194	40	22	538	271	50	24	458	308	67	23	388	313	81	19	267	226	85	22.5
nor	421	267	63	2	461	281	61	9	538	348	65	13.5	460	311	68	22	398	307	77	22	267	251	94	12

TABLE IV

Second Preferred List

Grade	3d Year				4th Year				5th Year				6th Year				7th Year				8th Year			
	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank	No. taking part	No. correct	%	Rank
already	371	60	16	15	486	205	42	12	437	186	43	19	364	225	62	19	352	154	44	23	257	199	77	22
beginning	371	33	9	23	486	119	25	24	437	162	37	23.5	364	166	46	23	352	231	66	21	257	192	75	23
chicken	371	183	49	1	487	341	70	1.5	436	362	83	1	365	329	90	1	352	338	96	1	257	256	99	1
choose	371	83	22	9	463	157	34	18	436	209	48	17.5	365	219	60	20	348	227	65	22	257	210	82	19.5
circus	371	75	20	11.5	463	180	39	14	438	219	50	15	366	264	72	14	351	262	75	17	257	245	95	6
grease	371	39	11	21	463	82	18	25	436	136	37	23.5	365	128	35	24	348	146	42	24	257	146	57	24
pigeons	371	24	7	25	463	134	29	22	436	177	41	20	365	208	57	21	348	244	70	20	257	210	82	19.5
quarrel	371	54	15	16.5	484	191	39	14	430	229	53	14	362	272	75	10.5	350	300	86	9	257	241	94	8.5
saucy	371	53	14	18.5	463	161	35	16.5	438	174	40	21	366	191	52	22	351	248	71	19	257	200	78	21
tailor	371	141	38	4	486	267	55	5	437	306	70	3	364	273	75	10.5	342	277	81	13	257	217	84	17.5
telegram	371	57	15	16.5	484	149	31	20	430	168	39	22	362	228	63	18	350	254	73	18	257	216	84	17.5
telephone	371	30	8	24	484	167	35	16.5	430	207	48	17.5	362	244	67	16.5	350	239	83	12	257	223	87	14.5
tobacco	371	43	12	20	463	179	39	14	436	262	60	12	365	274	75	10.5	348	305	88	6	257	248	96	4.5
too	371	52	14	18.5	487	134	28	23	436	117	27	25	365	86	24	25	349	104	30	25	257	111	43	25
towel	371	88	24	8	463	205	44	10	436	280	64	9.5	365	267	73	13	348	271	78	14	257	241	94	8.5
Tuesday	371	169	46	2	487	340	70	1.5	436	291	67	7	365	292	80	7	349	303	87	7.5	257	233	91	11.5
tying	371	162	44	3	487	282	58	4	436	305	70	3	362	246	68	15	350	266	76	16	257	224	87	14.5
whole	421	71	17	14	457	207	43	11	537	346	64	9.5	460	358	78	8	395	333	84	10.5	267	239	90	13
against	421	80	19	13	480	146	30	21	538	291	54	13	458	345	75	10.5	395	332	84	10.5	267	251	94	8.5
answer	426	116	27	7	486	228	47	9	536	360	67	7	453	389	86	2.5	395	360	90	4.5	267	259	97	2.5
butcher	424	140	33	5	487	288	59	3	536	372	69	5	453	387	85	4.5	395	356	90	4.5	267	259	97	2.5
guess	425	83	20	11.5	462	147	32	19	537	262	49	16	459	309	67	16.5	393	302	77	15	267	228	85	16
instead	421	134	32	6	484	231	48	8	535	332	62	11	459	393	86	2.5	396	346	87	7.5	267	242	91	11.5
raise	426	90	21	10	485	261	54	6	535	356	67	7	459	387	84	6	396	369	93	3	267	252	94	8.5
beautiful	371	36	10	22	493	256	52	7	436	307	70	3	365	311	85	4.5	349	327	94	2	257	248	96	4.5

and hardest in the middle grades. "Whose" is a type of such a word. It presents no great difficulty until children learn the use of the apostrophe. Then they write "who's." Later they partly recover from this practice. There are quantities of words which show this dip in the middle grades. One homonym is often easy until the other has been consciously related to it. Analogies falsely assumed play a harmful rôle. The rapid enriching of the vocabulary as new subjects and new phases of old subjects are taken up in the middle grades probably induces some confusion. To just what extent this is true we do not know but we are sure that it is true to a significant degree.

Finally, we have the large class of words which—when a sufficient number of children are tested—do show an increase in correctness from grade to grade, but which do not advance in anything like a constant ratio to the advance from grade to grade in spelling ability. Even among the words chosen as most favorable this discrepancy may be seen. The word "sure" (Table III) is 47 per cent correct in the 3d grade. This gives it a rank of 6th among the 25 words of the First Preferred List for that grade. In the 4th grade it advances to 55 per cent correct, but this advance is not sufficient to maintain its position. It falls to a rank where it is tied with "whistle" for 11th and 12th place—i.e., its rank is 11.5. In the 5th, 6th, 7th, and 8th grades its rank remains fairly constant with the 4th-grade rank. It is 10.5, 14.5, 10.5, and 12.

If this sort of irregularity is found in a word chosen as among the most regular, it will easily be seen how much greater would be the irregularity of words which were rejected.

§ 6. *Examination of the First Preferred List*

To further establish the value of the lists, we may investigate the behavior of the words as between grades in each school and in all schools combined. We shall begin with the First Preferred List. This list was chosen in the first instance from the returns of School I. It was taken to be the 25 words of closest correlation between the grades of that school. It was then referred to the other schools and the correlations were worked out for them.

The method used was that suggested by Spearman in his article: "'Foot-rule' for Measuring Correlation" (Spearman, '06). This method is explained and criticised by Brown in "The Essentials of Mental Measurement," pp. 71-76 (Brown, '11), and in a more elementary way by Whipple in his "Manual," pp. 34 and 35. (Whipple, '10.)

The formula is

$$R = 1 - \frac{\Sigma(g)}{1/6(N^2 - 1)}$$

where $\Sigma(g)$ denotes the sum of the "gains" in rank (sum of positive differences) of the second series on the first, and $1/6$ $(N^2 - 1)$ is the value of the sum of such gains which may be expected by chance. These R-values were then expressed as r-values (Pearson Coefficients of Correlation) by means of a table of equivalents. This table (Whipple, '10, p. 36), has been worked out from Spearman's conversion formula $r = \sin\left(\frac{\pi}{2} R\right)$.

The method is illustrated for the 4th and 5th grades of School I in Table V.

By the formula, $R = 1 - \frac{\Sigma(g)}{1/6\,(N^2 - 1)}$, $R = .72$. This is equivalent to an r-value of .90. The correlation is therefore very satisfactory. Its Probable Error is .026, which is so small in relation to the obtained correlation, that the latter has a very high degree of reliability.

The correlation of each grade with every other grade for School I was as follows:

Correlation of	4th	grade with	5th	grade,	.90
"	"	4th	"	"	6th	"	.88
"	"	4th	"	"	7th	"	.88
"	"	5th	"	"	6th	"	.92
"	"	5th	"	"	7th	"	.93
"	"	6th	"	"	7th	"	.88
				Average.....	.90
				(P. E. = .026)	

TABLE V

SCHOOL I. COEFFICIENTS OF CORRELATION FOR 4TH AND 5TH GRADES DERIVED. FOOT-RULE METHOD

	4th Grade		5th Grade		Gains, 5th on 4th
	%	Rank	%	Rank	
even	90	1	94	2	1
lesson	80	2	85	4.5	2.5
only	77	3.5	96	1	...
smoke	77	3.5	81	7	3.5
front	70	5.5	75	11	5.5
sure	70	5.5	85	4.5	...
pear	69	7	77	10	3
bought	67	8	83	6	...
another	65	9	79	8.5	...
forty	63	10.5	72	13	2.5
pretty	63	10.5	89	3	...
wear	56	12	73	12	...
button	50	13	70	14	1
minute	49	14	65	16	2
cousin	42	15.5	65	16	.5
nails	42	15.5	79	8.5	...
janitor	40	17.5	56	19	1.5
saucer	40	17.5	65	16	...
stopping	30	19	47	23	4
sword	25	20	50	22	2
freeze	21	22	63	18	...
touch	21	22	52	21	...
whistle	21	22	55	20	...
carriage	17	24	45	24	...
nor	13	25	43	25	...
					$29 = \Sigma(g)$

It may be interesting before taking up the results in other schools to see to what extent the 'foot-rule' method justifies itself in this kind of work. The 'product-moment' ($r = \frac{\Sigma xy}{n\sigma_1\sigma_2}$) and the 'unlike signs' methods ($r = \cos \pi v$) were used as a check. Table VI shows the result.

It is evident from Table VI that if the true numerical statement of the amount of correlation may be expected to be at or near the average of all three methods, the one which tends most nearly to approximate the true result is here the 'foot-rule'

TABLE VI

r-VALUES BETWEEN GRADES OF SCHOOL I AS FOUND BY THREE METHODS

Pairs of Grades	Foot-rule	Product-moment	Unlike signs	Average
4th with 5th.....	.90	.92	.99	.94
4th " 6th.....	.88	.76	.97	.87
4th " 7th.....	.88	.73	.81	.81
5th " 6th.....	.92	.78	.99	.90
5th " 7th.....	.93	.79	.88	.87
6th " 7th.....	.88	.91	.93	.91
Averages.....	.90	.82	.93	.88

method. It has therefore seemed justifiable in future computations of this sort to save the threefold labor and to rely upon this method.

It is to be expected that since the list of words was in the main selected from the results in School I, the correlation will prove to be higher in that school than in any other. Such is indeed the case as will be seen by an inspection of Table VII. It will be remembered that School I has no 8th grade and that the 3d grade in that school was not tested.

We find that the correlations in School II are considerably lower than in schools III, IV, and V, even showing an apparent inverse relation in one instance. Yet the average of the coefficients for School II is .42, the P.E. of which is less than .11. A correlation is entitled to scientific consideration if it is more than twice as large as its probable error. This one is nearly four times its probable error, and may therefore be regarded as satisfactory. Still more so are the relations between grades in the other schools; while, for all schools combined, the coefficients of grade-to-grade correlation range from .47 to .93 with an average of .76, A.D =.12. Since for these same values P.E. ranges from .10 to .02 with an average at .057, the reliability of these values is adequate.

It appears therefore that this list of words possesses the advantage of maintaining practically the same order of difficulty throughout the grades from the 3d to the 8th. In any grade the hardest word, the easiest word, and the words which take rank between tend strongly to hold their positions in every

other grade. Our list then is to a marked degree independent of fluctuations between grades.

But this might be true and still leave something to be desired. Schools differ in many respects—in quality of teaching and supervision, in preferred methods, in word lists studied, in the character of the children as to economic and racial condition. The schools which we have under consideration differ widely in all those respects. These variations in local conditions may very likely produce considerable variation in the quality of the spelling output.

TABLE VII

COEFFICIENTS OF CORRELATION. GRADE WITH GRADE, AND EACH GRADE WITH ALL GRADES FOR EACH SCHOOL. FIRST PREFERRED LIST

School	II	III	IV	V	All schools
3d with 4th	.37	.78	.67	.69	.79
" " 5th	.25	.55	.31	.71	.71
" " 6th	—.07	.40	.45	.75	.55
" " 7th	.40	.23	.34	.75	.47
" " 8th	.01	.31	.47	.67	.71
" " entire school	*.35*	*.79*	*.77*	*.85*	*.82**
4th " 5th	.81	.59	.61	.89	.83
" " 6th	.54	.59	.62	.84	.69
" " 7th	.37	.37	.60	.73	.62
" " 8th	.20	.34	.62	.77	.72
" " entire school	*.78*	*.84*	*.84*	*92*	*.90**
5th " 6th	.72	.60	.90	.93	.90
" " 7th	.60	.69	.88	.90	.86
" " 8th	.52	.48	.83	.72	.93
" " entire school	*.91*	*.84*	*.82*	*.95*	*.98**
6th " 7th	.56	.66	.94	.89	.90
" " 8th	.62	.60	.85	.89	.89
" " entire school	*.77*	*.77*	*.84*	*.93*	*.88**
7th " 8th	.45	.76	.90	.80	.89
" " entire school	*.75*	*.61*	*.78*	*.89*	*.86**
8th " entire school	*.60*	*.57*	*.82*	*.81*	*.91**
Average, grade with grade	.42	.53	.67	.80	.76
Average, grade with school	*.69*	*.74*	*.81*	*.89*	*.89**

* These *r*-values are for each grade with the grades of all schools

TABLE VIII

CORRELATIONS OF SCHOOL WITH SCHOOL AND OF EACH SCHOOL WITH ALL SCHOOLS FOR EACH GRADE. FIRST PREFERRED LIST

School I is not included because of its different system of grading

Grades	3d grade	4th grade	5th grade	6th grade	7th grade	8th grade	All grades
School II with III	.45	.60	.82	.55	.51	.31	.70
" II " IV	.32	.74	.66	.78	.56	.56	.80
" II " V	.61	.82	.87	.60	.59	.29	.88
" II " All	.62	.76	.93	.87	.88	.61	.91*
" III " IV	.76	.66	.60	.73	.62	.60	88
" III " V	.54	.77	.79	.78	.47	.50	.81
" III " All	.88	.89	.87	.73	.77	.61	.87*
" IV " V	.77	.75	.78	.67	.60	.64	.85
" IV " All	.83	.84	.82	.93	.83	.83	.93*
" V " All	.83	.86	.95	.78	.79	.85	.93*
Average, school with school	.58	.72	.75	.69	.56	.48	.82
Average, each school with all schools	.79	.84	89	.83	.82	.73	.91*

* These figures are for each school with all grades and schools combined, i.e., with all participants.

A method in reading and word study which makes extensive use of the phonogram may possibly cause some words to become easy which are otherwise difficult. If the pupils in one school come from homes where English is not spoken, they may find difficult a set of words other than that which children of English-speaking parents find difficult.

With the purpose of throwing some light on this point we shall consider what the correlation is *between schools* for each grade and for all grades with respect to the First Preferred List. Table VIII shows the correlation coefficients. The school-with-school average correlations range from .48 to .82 with a median at .69. The school-with-all-school averages range from .73 to .91 with a median at .83. A few of the coefficients throughout the table are low. There are, however, but six that are below .50. All but one of these are in the extreme grades (3d or 8th) and have to do with School II. The circumstances under which this school was examined account for this. The tests were given

immediately after the long summer vacation and the test-material comprised the Original List (270 words). The other schools were tested considerably later in the school year and the pupils in those schools wrote the Selected List (100 words).

Notwithstanding these few shortcomings the 70 coefficients of Table VIII form an impressive argument for the value of the First Preferred List. We may fairly contend that not only are the positions of the words of this list relatively stable as between grades (Table VII), but that this permanency holds as between schools.

§ 7. *Examination of the Second Preferred List*

The second list of 25 words (see p. 13 or Appendix II) was examined in the same way that the first list was examined, i.e., with reference to correlations first between grades, and second between schools.

At School I the correlations between grades were found to be as follows (Compare with similar tabulation for the First Preferred List on page 17:

4th	grade	with	5th	grade	.87
4th	"	"	6th	"	.83
4th	"	"	7th	"	.79
5th	"	"	6th	"	.95
5th	"	"	7th	"	.78
6th	"	"	7th	"	.83
			Average....		.84
			(P.E.=.04)		

For the other schools Table IX shows the correlations. It may be compared with Table VII (page 20).

A comparison of Table IX with Table VII shows that although the word list to which Table IX refers was taken, on the basis of partial knowledge, to be somewhat inferior to the First Preferred List, these coefficients do not show it. The grade-with-grade averages (.66, .67, .60, .48, and .75) are higher for some schools than in Table VII and lower for others. Their central tendency is almost identical while the closeness of grouping is greater for the second than for the first list. Of the

105 measures of correlation in Table IX only 13 are less than four times their probable error, and only 3 are less than twice their probable error. The grade-to-grade relationships for all schools (column 6) range from .40 to .95, average .75, A.D.= .14. This is satisfactory to a degree scarcely, if at all, less than is the showing for the first list.

TABLE IX

COEFFICIENTS OF CORRELATION. GRADE WITH GRADE AND EACH GRADE WITH ALL GRADES FOR EACH SCHOOL. SECOND PREFERRED LIST

School	II	III	IV	V	All schools
3d with 4th	.60	.75	.55	.38	.74
" " 5th	.69	.62	.56	.26	.73
" " 6th	.61	.55	.55	.11	.55
" " 7th	.43	.55	.41	.07	.40
" " 8th	.51	.52	.35	—.01	.41
" " entire school	*.73*	*.84*	*.75*	*.34*	*.72**
4th " 5th	.73	.82	.69	.48	.90
" " 6th	.74	.61	.44	.40	.80
" " 7th	.60	.62	.45	.38	.69
" " 8th	.57	.60	.51	.38	.62
" " entire school	*.82*	*.88*	*.78*	*.67*	*.93**
5th " 6th	.75	.76	.74	.84	.90
" " 7th	.55	.75	.73	.65	.83
" " 8th	.62	.73	.72	.83	.80
" " entire school	*.81*	*.92*	*.90*	*.92*	*.95**
6th " 7th	.83	.67	.79	.76	.94
" " 8th	.82	.71	.77	.84	.89
" " entire school	*.93*	*.82*	*.84*	*.88*	*.91**
7th " 8th	.80	.80	.74	.86	.90
" " entire school	*.88*	*.83*	*.80*	*.78*	*.87**
8th " entire school	*.89*	*.80*	*.81*	*.86*	*.80**
Average, grade with grade	.66	.67	.60	.48	.75
Average, grade with school	*.84*	*.87*	*.81*	*.74*	*.86**

* These *r*-values are for each grade with all grades of all schools.

Table X (whose counterpart for the first list is Table VIII) reveals, as Table IX did not, the relative inferiority of the second list. There are 49 coefficients in Table X that are lower than the corresponding figures in Table VIII. Only 21 are

higher. The school-with-school averages are lower in five instances and higher in but two. The order of difficulty of these words is therefore not so permanent as between schools. It is, however, sufficient abundantly to justify the list. There are but six of the 70 coefficients in the body of the table that are less than four times their probable error, and but three that are less than twice their probable error. In some respects, indeed, this list is superior to the first list. A comparison of

TABLE X

CORRELATIONS OF SCHOOL WITH SCHOOL AND OF EACH SCHOOL WITH ALL SCHOOLS FOR EACH GRADE. SECOND PREFERRED LIST

School I not included. Compare with Table VIII, p. 21

	3d grade	4th grade	5th grade	6th grade	7th grade	8th grade	All grades
School II with III	.48	.51	.47	.65	.60	.52	.66
" II " IV	.29	.17	.52	.65	.45	.38	.43
" II " V	.05	−.08	.52	.55	.43	.64	.60
" II " All	*.57*	*.56*	*.84*	*.95*	*.76*	*.83*	*.84**
" III " IV	.55	.38	.82	.60	.61	.67	.59
" III " V	.47	.43	.62	.59	.72	.84	.87
" III " All	*.76*	*.78*	*.78*	*.80*	*.85*	*.84*	*.88**
" IV " V	.47	.43	.55	.59	.81	.66	.65
" IV " All	*.79*	*.76*	*.74*	*.75*	*.65*	*.69*	*.77**
" V " All	*.62*	*.57*	*.82*	*.65*	*.84*	*.88*	*.78**
Average, school with school	.39	.31	.58	.61	.60	.62	.63
Average, each school with all schools	*.69*	*.68*	*.80*	*.79*	*.78*	*.81*	*.82**

* These figures are for each school with all grades and schools combined, i.e., with all participants.

Tables III and IV shows that the words of the second list are in general more difficult than those of the first. Doubtless the first list is a somewhat better test for lower grades, while the second is a better test for upper grades. This supposition is neatly supported by the figures in Table X. Sixteen of the 21 that are higher in this table than in Table VIII are in the three upper grades, while the two higher *average* correlations are in the 7th and 8th grades. If, therefore, some of the words in the first list are found to be too easy for the highest grades—as

doubtless they may be—then the second list will supply a valuable supplement to the first.

§ 8. *Conclusions Regarding the Preferred Lists*

Our lists therefore prove to be well selected. Success and failure in spelling them may be used with considerable confidence to measure the thing we call spelling ability. The establishment of this fact is of the utmost importance. In general when we are to measure mental traits or capacities the thing we directly measure is itself a physical phenomenon or fact. We measure fatigue by the number and height of lifts with the ergograph, or by the distance between points of the esthesiometer necessary to be recognized as 'two.' We measure attention and perception by counting dots or by cancellation; memory, by the number of digits reproduced; association by the number of words pronounced in a given time; and intelligence itself, by a series of tests which may be scored objectively. What we deal with directly is something, assumed to be functionally related to the trait in question, which can be measured in time or space or which can be counted. If this objective manifestation does not accurately register the subjective state to which it is supposed to correspond, it is impaired, to the extent of its inaccuracy, as an index to be directly measured.

Now it is undoubtedly true that the misspellings of most words are unreliable as indicating lack of spelling ability in general; and on the other hand it is probable that to spell them correctly often argues little more than that the subject can spell the particular words that he did spell. Most words are in some way special—and they are special (particularly for children) in ways that we do not realize. Very often they do not mean the same thing to one person that they do to another. They are frequently pronounced differently by different people. They suggest dissimilar imagery. They connote variously. They range from very easy to very hard; and those that are easy for some people are hard for others. Moreover there are numerous ways of misspelling them, each indicating its own causal incoördination. An error may not be equal to an error even in misspelling the same word.

It would be presumptuous to suppose that all these difficulties have been overcome in selecting our two preferred lists. Without doubt we have only roughly approximated the ideal conditions under which a physical fact may be the transcript of a mental trait. Probably nothing more satisfactory than an approximation can be devised. But we have been at no small pains to secure a list of words which would be free from many of these variations, and we think we have done so. From an inspection of them we may be justified in believing that their pronunciation and meaning are fairly constant for everybody; and we may regard it as probable that their associative connections do not vary much for different people. From a statistical analysis of them we find that their behavior with elementary school children is constant to a marked degree, and in particular that it is relatively independent of maturity and of local conditions. Older children in higher grades spell them more frequently and in each grade more frequently than in the one before it. Children in schools under favorable circumstances do better with them than do children in less favorable situations. It is because they reflect these conditions that they are valuable. By the use of them, conditions in a school, a class, or an individual may be revealed; and conversely to a certain extent if the conditions are known (e.g., the grade) the results, by the use of them, are predictable.

These lists, then, tend strongly to remain intact under various conditions. As lists they appear to be reliable, and our numerical results give quantitative expression to this reliability. But as to the words themselves, we shall yet have much to say. There has been no attempt to secure lists composed of words of equal difficulty. The effort has rather been to choose words which differ widely in this respect. We shall now attempt to arrange these words on a scale which shall accurately represent their difficulty, assuming as true a certain supposition concerning the form of distribution of spelling ability within a school grade. The resulting scale will represent their difficulty approximately in so far as this supposition is approximately true. When this is done, their value for test purposes independently of the list which contains them will be established.

§ 9. *Ratings of Individual Pupils*

In addition to scoring words, the papers of individual pupils were rated. This was done in the usual manner, the ability to spell one word being scored as equal to the ability to spell any other word of the list. This procedure is subject to the criticism made in Section 1 above; but in the absence of any evaluation of the words, a system of weighting is not possible. The results will not be misused here.

The test material consisted of the 100 words of the Selected List. The papers written at schools II, III, IV, and V were used. School I was not available because of its system of grading. A few papers were incomplete in each of the schools; these were rejected in this part of the work. In all, 2,487 papers were rated. Table XI shows for each grade the distribution of in-

TABLE XI

DISTRIBUTION OF INDIVIDUAL RATINGS OF PUPILS IN SCHOOLS II, III, IV AND V

Percentage Correct	3d Grade		4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
0– 5	9	2.0	1	.2	1	.2						
6– 10	22	4.9	1	.2	2	.4						
11– 15	30	6.7	10	2.1	1	.2						
16– 20	38	8.5	12	2.6	2	.4						
21– 25	44	9.9	13	2.8	6	1.2	2	.5				
26– 30	47	10.6	23	4.9	12	2.3						
31– 35	34	7.6	29	6.2	13	2.5	2	.5	2	.5		
36– 40	38	8.5	27	5.8	11	2.1						
41– 45	24	5.4	30	6.4	18	3.5	6	1.4	2	.5		
46– 50	34	7.6	33	7.1	28	5.4	4	1.0	1	.3		
51– 55	26	5.8	27	5.8	20	3.9	6	1.4	3	.8		
56– 60	24	5.4	31	6.6	32	6.2	15	3.6	5	1.4	1	.4
61– 65	26	5.8	39	8.4	44	8.5	12	2.9	6	1.6	1	.4
66– 70	17	3.8	29	6.2	48	9.3	23	5.5	8	2.2	3	1.1
71– 75	13	2.9	45	9.6	49	9.5	30	7.2	18	4.9	8	2.9
76– 80	8	1.8	35	7.5	59	11.5	52	12.4	31	8.5	11	4.0
81– 85	4	.9	33	7.1	37	7.2	67	16.0	38	10.4	19	6.9
86– 90	4	.9	26	5.6	64	12.4	61	14.6	79	21.6	41	14.8
91– 95	3	.7	19	4.1	50	9.7	101	24.2	93	25.5	80	28.9
96–100			4	.9	18	3.5	37	8.9	79	21.6	113	40.8
Totals...	445		467		515		418		365		277	
Medians.		35.8		60.70		73.10		84.90		90.50		94.68
A. D....		18.0		20.9		10.4		10.0		7.9		5.8

dividual ratings. It reads as follows: "In the 3d grade 9 pupils were rated between 0% and 5%, which was 2.0% of all the 3d grade pupils. In the 4th grade 1 pupil was rated between 0% and 5%, which was .2% of all the 4th grade pupils," etc.

The striking characteristic of the distribution of these ratings is their extreme variability. Children of the 3d grade are represented in every group but one, while children of the 4th and 5th grades are rated in every group. It appears that we may expect a few 6th- and 7th-grade children to spell not more than 20 or 30 of these hundred words, which is not quite as good as the typical ability of 3d-grade children for the same words. The 8th-grade pupils show the least variation. This is probably true of this grade in general. It is not, however, as marked as these figures indicate. The 100-word list as a whole, whatever may be true about some of the individual words, did not thoroughly test this grade. A glance at Fig. 6 will show how sharply cut off at the high end is the curve of distribution. This is against all the facts which we know about eighth-graders in particular and human ability in general. A harder test would have shown a lower mode and a more gradual tapering off at the upper end of the curve. But even as this record stands we may look for a considerable number—between 7 and 8 per cent—of 8th-grade pupils to average no better than the typical performance of 5th-grade children.

TABLE XII

DISTRIBUTION OF INDIVIDUAL RATINGS GROUPED TO SHOW MODES. FIGS. 1–7

Percentage Correct	3d Grade		4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
0– 10	6.9	22.1	.4	5.1	.6	1.2						
11– 20	15.2		4.7		.6							
21– 30	20.5	36.6	7.7	19.7	3.5	8.1	.5	1.0				
31– 40	16.1		12.0		4.6		.5		.5	.5		
41– 50	13.0	24.2	13.5	25.9	8.9	19.0	2.4	7.4	.8	3.0		.4
51– 60	11.2		12.4		10.1		5.0		2.2		.4	
61– 70	9.6	14.3	14.6	31.7	17.8	38.8	8.4	28.0	3.8	17.2	1.5	8.4
71– 80	4.7		17.1		21.0		19.6		13.4		6.9	
81– 90	1.8	2.5	12.7	17.7	19.6	32.8	30.6	63.7	32.0	79.1	21.7	91.4
91–100	.7		5.0		13.2		33.1		47.1		69.7	

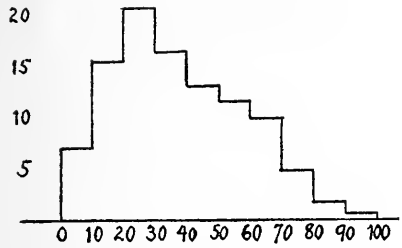


Fig. 1

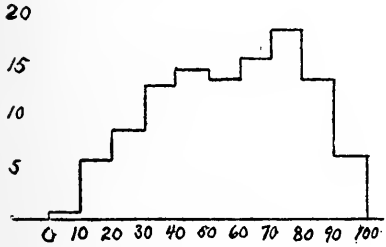


Fig. 2.

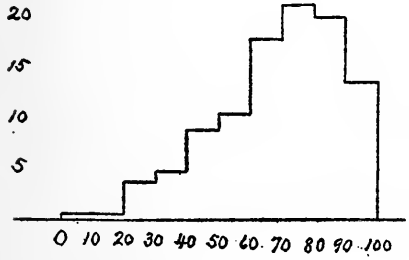


Fig. 3.

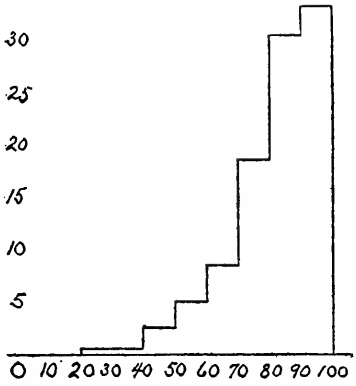


Fig. 4.

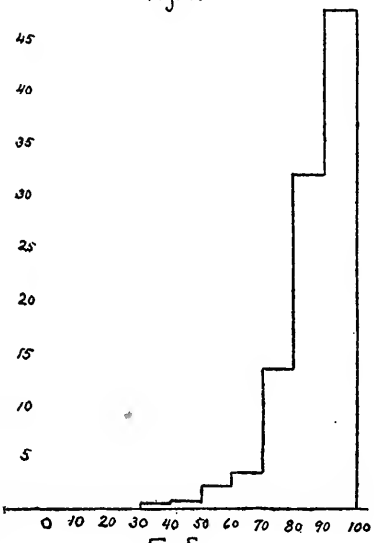


Fig. 5.

FIG. 1. Frequency of each rating (of per cent of Selected List spelled correctly) in grade 3.

FIG. 2. Frequency of each rating (of per cent of Selected List spelled correctly) in grade 4.

FIG. 3. Frequency of each rating (of per cent of Selected List spelled correctly) in grade 5.

FIG. 4. Frequency of each rating (of per cent of Selected List spelled correctly) in grade 6.

FIG. 5. Frequency of each rating (of per cent of Selected List spelled correctly) in grade 7.

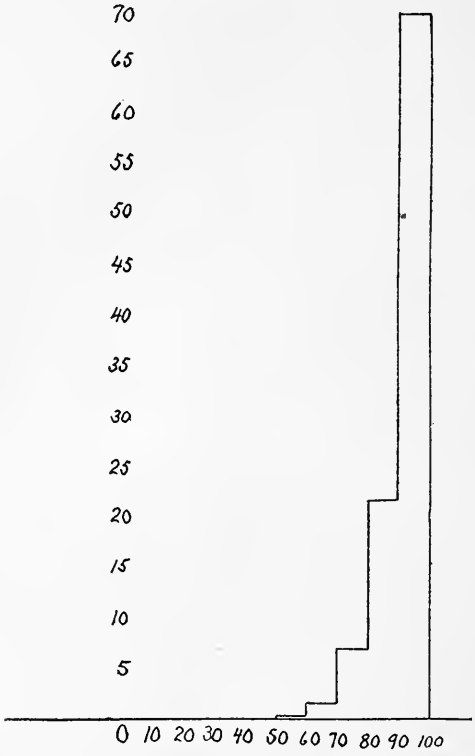


FIG. 6

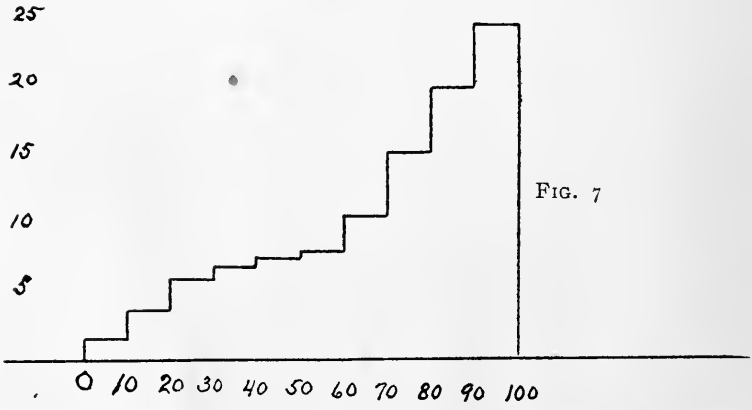


FIG. 7

FIG. 6. Frequency of each rating (of per cent of Selected List spelled correctly) in grade 8.

FIG. 7. Frequency of each rating (of per cent of Selected List spelled correctly) in grades 3-8 combined.

Another characteristic of the distributions shown in Table XI is the absence of clearly marked modes. Table XII is a grouping of the per cent columns of Table XI into 10's and 20's. From this grouping wide modes of marked character are shown.

Figs. 1 to 7 show the same facts graphically. From the nature of these figures the test material appears to have been capable of revealing satisfactorily the spelling ability of grades 3, 4, and 5. Figs. 1 to 7 are the surfaces of frequency of spelling ability with the Selected List. In each of them the horizontal scale shows percentages correct; the vertical scale shows the per cent of children having each rating for percentage correct, by steps of 10. The number of children represented is 445 in grade 3, 467 in grade 4, 515 in grade 5, 418 in grade 6, 365 in grade 7, and 277 in grade 8.

§ 10. *Overlapping*

It follows as a matter of course from the variability of these ratings that the overlapping of grade on grade is conspicuous. We have located the median abilities, of each grade, for the selected word list. They are: 3d grade, 35.8; 4th grade, 60.7; 5th grade, 73.1; 6th grade, 84.9; 7th grade, 90.5; 8th grade, 94.7 (See Table XI). Table XIII shows the number of pupils and the per cent of pupils in each grade whose ratings equalled or exceeded the medians of every other grade. The table reads as follows: In the 3d grade 76 pupils equalled or exceeded the median rating of the 4th grade which was 17.1% of all the 3d-grade pupils; 27 equalled or exceeded the median rating of the 5th grade which was 6.1% of all the 3d-grade pupils. In the 4th grade 378 pupils equalled or exceeded the median rating of the 3d grade which was 80.9% of all the pupils of the 4th grade, etc. It will be noticed that there are two places where there is no overlapping. There are no 3d-grade children who equal the median rating of the 8th grade, although the 3 who exceed the 7th-grade median come very near it. Two of them are rated at 93 and one at 94, only 1.7 and .7 below the 8th-grade median. There is also no overlapping of the 8th grade on the 3d. All the pupils of the 8th grade exceed the median of the 3d grade. When, however, we say that at these points there is no overlapping, we do not mean that their surfaces of frequency do not enclose

TABLE XIII

NUMBER AND PER CENT OF PUPILS IN EACH GRADE WHOSE ABILITY EQUALLED OR EXCEEDED THAT OF THE MEDIAN PUPIL IN EVERY OTHER GRADE

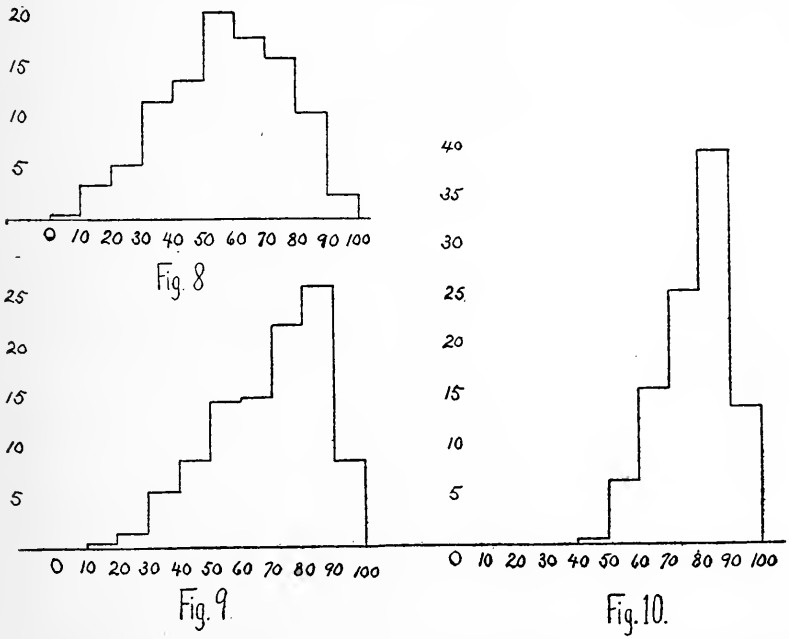
		3d Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade
3d grade.....	No.		76	27	9	3	0
N=445.....	%		17.1	6.1	2.0	0.7	0
Med.=35.8							
4th grade.....	No.	378		146	52	27	9
N=467.....	%	80.9		31.3	11.1	5.8	1.9
Med.=60.7							
5th grade.....	No.	478	370		142	73	30
N=515.....	%	92.8	71.8		27.6	14.2	5.8
Med.=73.1							
6th grade.....	No.	414	384	338		142	57
N=418.....	%	99.0	91.9	80.1		34.0	13.6
Med.=84.9							
7th grade.....	No.	363	354	328	256		99
N=365.....	%	99.5	96.4	89.9	70.1		27.1
Med.=90.5							
8th grade.....	No.	277	276	269	241	200	
N=277.....	%	100	99.6	97.1	87.0	72.2	
Med.=94.7							

common areas. If Fig. 1 is placed on Fig. 8 so that the zero points coincide, it is evident that there is considerable area common to both. We mean that the upper part of the 3d-grade surface of frequency does not lap over *the median point* of the 8th-grade surface, and that the lower part of the 8th-grade surface does not reach down to the 3d-grade median. There are many 3d-grade children who do better than the poorest 8th-grade children.

The fact is, then, that except as between the 3d and 8th grades, some pupils of each grade perform like typical children of every other grade. Since this is true, it serves to fix the location of the frequency curves and medians for each grade with reference to each other. For the purpose of doing so we shall for the present assume that the distribution of spelling ability in each grade is "normal," i.e., is correctly represented by the curve of error.

There is some argument for this assumption. The fact that

our surfaces of frequency (Figs. 1-6) do not, especially for upper grades, closely resemble the normal curve, only shows that the test material was not difficult enough to bring out a distribution in real accordance with spelling ability. The result of using a different list of words is shown for grades 6, 7, and 8 by Figs. 8, 9, and 10. The test material in this instance was Rice's "Sentence Test": 396 children in the 6th grade, 367 in the 7th, and 244 in the 8th wrote this test in schools II, III, and VIII. The



FIGS. 8, 9 and 10. Frequency of each rating (of per cent of Rice Sentence List spelled correctly) in grades 6, 7 and 8, respectively. N=396 for grade 6; 367 for grade 7; and 244 for grade 8. The horizontal scale is for per cent spelled correctly; the vertical scale is for the percentage of children receiving each rating for percentage correct, by steps of 10.

surface of frequency for the 6th grade is close to the "normal" surface. If that for the 7th and 8th grades is less so, it is still far more regular than the surfaces shown for these grades in Figs. 5 and 6 and might be made still more so by an appropriate selection of test material. There seems no good ground for as-

suming that the distribution of spelling ability in any grade is not according to the normal curve or according to a curve which resembles it closely.

§ 11. *Location of Grade Medians*

Upon the assumption, therefore, of a normal distribution we may proceed to locate the grade medians with reference to each other. In all cases we shall work with per cents instead of with numbers of pupils. This will reduce all surfaces of frequency to equal areas. We shall assume further that the real variability of any one of these grades in spelling ability is equal to the real variability of any other one of them.

We have already seen (Table XIII) that 17.1% of the 3d-

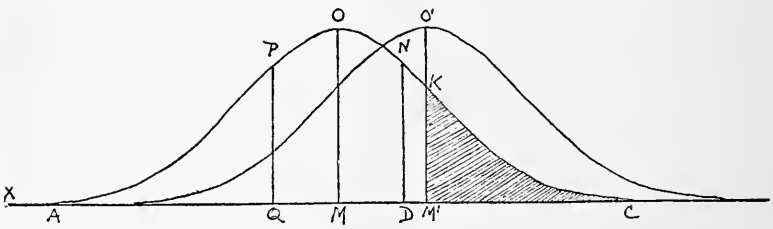


FIG. 11. Showing the overlapping of the 3d and 4th grade surfaces of frequency.

grade pupils equal or exceed the median ability of 4th-grade children. Fig. 11 shows this fact by a diagram. The surface on the left (Axis OM) represents the 3d-grade distributions. M is its median point, MD $(=MQ)$ is its probable error—i.e., figure $NPQD$ is one-half its area, thus graphically representing one-half the cases in the 3d grade, which accordingly do not deviate from the median by an amount greater than MD. The surface on the right represents the 4th-grade distribution beyond whose median axis, O^1M^1, the 3d-grade surface extends to an amount represented by the shaded figure KCM^1. This stands for the 3d-grade children who equal or exceed the 4th-grade median—i.e., it is 17.1% of the 3d-grade surface of frequency. Accordingly the area OKM^1M represents 32.9% of the cases. The distance MM^1 may now be found in terms of P.E. It is the distance from the median point along the X-axis of the normal probability integral which includes 32.9% of the cases. Distances

corresponding to different per cents of the total area of the curve have been worked out. Table XIV, which is taken from

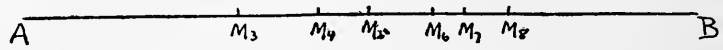
TABLE XIV

TABLE OF VALUES OF THE NORMAL PROBABILITY INTEGRAL CORRESPONDING TO VALUES OF P.E. TOTAL AREA OF THE SURFACE OF FREQUENCY TAKEN AS 10,000

X / P.E.	No. Cases	Δ	X / P.E.	No. Cases	Δ	X / P.E.	No. Cases	Δ	X / P.E.	No. Cases	Δ
0	0		1.5	3441		3.00	4785		4.5	4988	
		135			81			18			1
.05	135		1.55	3521		3.05	4802		4.55	4989	
		134			80			17			1
.1	269		1.6	3597		3.1	4817		4.6	4990	
		134			76			15			1
.15	403		1.65	3671		3.15	4831		4.65	4991	
		133			74			14			1
.2	536		1.7	3742		3.2	4845		4.7	4992	
		134			71			14			1
.25	670		1.75	3811		3.25	4858		4.75	4993	
		132			69			13			1
.3	802		1.8	3876		3.3	4870		4.8	4994	
		131			65			12			.6
.35	933		1.85	3939		3.35	4881		4.85	4994.6	
		130			63			11			.6
.4	1063		1.9	4000		3.4	4891		4.9	4995.2	
		130			61			10			.5
.45	1193		1.95	4057		3.45	4900		4.95	4995.7	
		128			57			9			.5
.5	1321		2.0	4113		3.5	4909		5.0	4996.2	
		126			56			9			.4
.55	1447		2.05	4166		3.55	4917		5.05	4996.6	
		124			53			8			.5
.6	1571		2.1	4217		3.6	4924		5.1	4997.1	
		124			51			7			.3
.65	1695		2.15	4265		3.65	4931		5.15	4997.4	
		121			48			7			.3
.7	1816		2.2	4311		3.7	4937		5.2	4997.7	
		119			46			6			.3
.75	1935		2.25	4354		3.75	4943		5.25	4998.0	
		118			43			6			.2
.8	2053		2.3	4396		3.8	4948		5.3	4998.2	
		115			42			5			.2
.85	2168		2.35	4435		3.85	4953		5.35	4998.4	
		113			39			5			.2
.9	2281		2.4	4472		3.9	4957		5.4	4998.6	
		111			37			4			.2
.95	2392		2.45	4508		3.95	4961		5.45	4998.8	
		108			36			4			.2
1.0	2500		2.5	4541		4.0	4965		5.5	4999.0	
		106			33			4			.1
1.05	2606		2.55	4573		4.05	4968		5.55	4999.1	
		103			32			3			.1
1.1	2709		2.6	4602		4.1	4971		5.6	4999.2	
		101			29			3			.1
1.15	2810		2.65	4631		4.15	4974		5.65	4999.3	
		98			29			3			.1
1.2	2908		2.7	4657		4.2	4977		5.7	4999.4	
		96			26			3			.1
1.25	3004		2.75	4682		4.25	4979		5.75	4999.5	
		93			25			2			.05
1.3	3097		2.8	4705		4.3	4981		5.8	4999.55	
		91			23			2			.05
1.35	3188		2.85	4727		4.35	4983		5.85	4999.6	
		87			22			2			.05
1.4	3275		2.9	4748		4.4	4985		5.9	4999.65	
		85			21			2			.05
1.45	3360		2.95	4767		4.45	4987		5.95	4999.7	
					19						

Thorndike ('13, p. 200), presents these distances in units of P.E. By reference to it we find that 32.9% corresponds to 1.4088 P.E.

In a similar way, the 6.1% of 3d-grade children who equal or exceed the 5th-grade median (Table XIII) serve to locate that median at 2.2929 P.E. above the 3d-grade median. The 6th-grade median is 3.0441 P.E. and the 7th, 3.6429 P.E. above the 3d-grade median. The distance between the 3d- and 8th-grade medians cannot be directly calculated owing to the absence of sufficient overlapping.



Suppose the line AB to represent the range of spelling ability in the elementary school. At a certain distance above A, the absolute zero-point, stands the 3d-grade median, M_3. Above it and at distances to be determined are the medians of the 4th to the 8th grades, M_4———M_8. For brevity we shall call the distance from the 3d- to the 4th-grade median M_{3-4}, etc. M_{4-3} means the same distance as M_{3-4}, but measured in the opposite or negative direction.

TABLE XV

THE PER CENT OF PUPILS IN EACH GRADE WHOSE ABILITY EQUALLED OR EXCEEDED THAT OF THE MEDIAN PUPIL IN EVERY OTHER GRADE; WITH THE P.E. VALUES CORRESPONDING TO EACH PER CENT

		3d Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade
3d grade....	%		17.1	6.1	2.0	0.7	0
	P.E.		1.4088	2.2929	3.0441	3.6429	?
4th grade...	%	80.9		31.3	11.1	5.8	1.9
	P.E.	—1.2962		.7227	1.8111	2.3308	3.0767
5th grade ..	%	92.8	71.8		27.6	14.2	5.8
	P.E.	—2.1663	—.8553		.8819	1.5888	2.3308
6th grade...	%	99.0	91.9	80.1		34.0	13.6
	P.E.	—3.4500	—2.0735	—1.2532		.6117	1.6291
7th grade ..	%	99.5	96.4	89.9	70.1		27.1
	P.E.	—3.8200	—2.6673	—1.8918	—.7818		.9041
8th grade...	%	100	99.6	97.1	87.0	72.2	
	P.E.	?	—3.9375	—2.8114	—1.6704	—.8730	

Table XV gives all the distances between medians which our data permit us to calculate directly. The P.E. values, reading across the table, indicate that on the record of the pupils tested the 4th-grade median is found to be 1.4088 P.E. above the 3d-grade median, the 5th 2.2929 P.E. above it, the 6th 3.0441 P.E. above it, and the 7th 3.6429 P.E. above it; that the 3d-grade median is 1.2962 P.E. *below* the 4th-grade median, the 5th .7227 P.E. above it, etc.

It will be seen that M_4 is given as 1.4088 above M_3, while M_3 is given as only 1.2962 below M_4. We shall have to adopt one or the other, or some value between them as the most probably correct distance, M_{3-4}. Similarly for each of the other distances (except M_{3-8}) we have two values, and these two values are in each case somewhat different one from the other. The following are the pairs of values which Table XV shows:

M_{3-4}	1.4088	and	1.2962
M_{3-5}	2.2929	"	2.1663
M_{3-6}	3.0441	"	3.4500
M_{3-7}	3.6429	"	3.8200
M_{4-5}	.7227	"	.8553
M_{4-6}	1.8111	"	2.0735
M_{4-7}	2.3308	"	2.6673
M_{4-8}	3.0767	"	3.9375
M_{5-6}	.8819	"	1.2532
M_{5-7}	1.5888	"	1.8918
M_{5-8}	2.3308	"	2.8114
M_{6-7}	.6117	"	.7818
M_{6-8}	1.6291	"	1.6704
M_{7-8}	.9041	"	.8730

The differences between these pairs of values is in most cases small. In all cases they afford data for the determination of the distances between medians which will be probably more accurate than either of them.

We do not, however, need all these values. If we have five, namely, M_{3-4}, M_{4-5}, M_{5-6}, M_{6-7}, and M_{7-8}, all the others may be obtained by adding these together. We shall therefore attempt to derive as accurately as possible these five values in terms of the unit, P.E. Each of them is represented directly by two quantities as shown above. But it is clear that if we use more of the data of Table XV we may obtain values whose

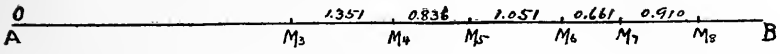
accuracy will be much more satisfactory. We may, for instance, find for the distance between the 4th-grade median and the 5th-grade median (M_{4-5}) a third value by subtracting from the distance between the 3d- and 5th-grade medians ($M_{3-5} = 2.2929$) the distance between the 3d- and 4th-grade medians ($M_{3-4} = 1.4088$). This gives .8841. Another value is the difference between the same two distances expressed negatively, i.e., according to our notation, between M_{5-3} (2.1663) and M_{4-3} (1.2962), which is .8701. Again we may use the difference between M_{4-6} and M_{5-6}, between M_{4-7} and M_{5-7}, between M_{4-8} and M_{5-8}; and for each of these differences between positive quantities we have a difference between corresponding negative quantities. This adds six more expressions, making ten altogether, for the same distance, M_{4-5}. This is of course only a beginning of the great number of combinations which may be used to get expressions for the same distance. We think these few, however, since they use each of the 18 segments (nine counted both ways) which terminate at either M_4 or M_5 will be sufficient to determine M_{4-5} with considerable accuracy. We doubt whether the remoter segments (e.g., M_{6-7}, M_{6-8}, and M_{7-8}) would, if used, increase the accuracy at all.

Accordingly we have calculated 10 values for M_{4-5}, M_{5-6}, and M_{6-7}. Since we have no expression of direct relation between M_3 and M_8, we have but 8 values for M_{3-4} and M_{7-8}. Table XVI gives all these values with the derivation of each. It also gives the averages, unweighted and weighted, of the values for each of the five median intervals.

It was felt that to give each of these items the same weight was to fail to take account of their reliability. The direct values are, no doubt, most to be depended upon. Those computed by using a distance which passes over one median are less so. Those involving two or more of these "skips" are still less so and diminish in reliability as the number of "skips" increases. It will be found that in column 2 of Table XVI the first quantity .8841 is derived by using a value that involves one skip. M_{3-5} skips over M_4, while M_{3-4}, which is taken from it, presents no skips. The second quantity, .7227, is direct, and there are no skips. .9292 involves one skip, .7420 three skips (M_{4-7} skips M_5 and M_6, and M_{5-7} skips M_6), etc. It will be found upon trial

TABLE XVI

DIRECT AND DERIVED VALUES OF MEDIAN DISTANCES IN TERMS OF P.E.



	M_{3-4}	M_{4-5}	M_{5-6}	M_{6-7}	M_{7-8}
	1.4088 (direct)	.8841 (M_{3-5}—M_{3-4})	.7512 (M_{3-6}—M_{3-5})	.5988 (M_{3-7}—M_{3-6})	? (M_{3-8}—M_{3-7})
	1.5704 (M_{3-5}—M_{4-5})	.7227 (direct)	1.0884 (M_{4-6}—M_{4-5})	.5199 (M_{4-7}—M_{4-6})	.7459 (M_{4-8}—M_{4-7})
	1.2330 (M_{3-6}—M_{4-6})	.9292 (M_{4-6}—M_{5-6})	.8819 (direct)	.7069 (M_{5-7}—M_{5-6})	.7420 (M_{5-8}—M_{5-7})
	1.3121 (M_{3-7}—M_{4-7})	.7420 (M_{4-7}—M_{5-7})	.9771 (M_{5-7}—M_{6-7})	.6117 (direct)	1.0174 (M_{6-8}—M_{6-7})
	? (M_{3-8}—M_{4-8})	.7459 (M_{4-8}—M_{5-8})	.7017 (M_{5-8}—M_{6-8})	.7250 (M_{6-8}—M_{7-8})	.9041 (direct)
	1.2962 (direct)	.8701 (M_{5-3}—M_{4-3})	1.2837 (M_{6-3}—M_{5-3})	.3700 (M_{7-3}—M_{6-3})	? (M_{8-3}—M_{7-3})
	1.3110 (M_{5-3}—M_{5-4})	.8553 (direct)	1.2182 (M_{6-4}—M_{5-4})	.5938 (M_{7-4}—M_{6-4})	1.2702 (M_{8-4}—M_{7-4})
	1.3765 (M_{6-3}—M_{6-4})	.8203 (M_{6-4}—M_{6-5})	1.2532 (direct)	.6386 (M_{7-5}—M_{6-5})	.9196 (M_{8-5}—M_{7-5})
	1.1527 (M_{7-3}—M_{7-4})	.7755 (M_{7-4}—M_{7-5})	1.1100 (M_{7-5}—M_{7-6})	.7818 (direct)	.8886 (M_{8-6}—M_{7-6})
	? (M_{8-3}—M_{8-4})	1.1261 (M_{8-4}—M_{8-5})	1.1410 (M_{8-5}—M_{8-6})	.7974 (M_{8-6}—M_{8-7})	.8730 (direct)
Average	1.3326	.8471	1.0406	.6344	.9201
Weighted Average	1.3505	.8363	1.0505	.6608	.9101

that in all values there are either 0, 1, 3, or 5 skips. We have weighted them 6, 4, 3, and 2 respectively (ratio about 1.5). This is, of course, pure assumption, nor do we know of any convenient plan of weighting which would not be. All we can

say is that to us the direct values seem to be quite one and one-half times as reliable as those involving a distance which passes over one median, that it seems reasonable the latter should be about as many times more reliable than those involving 3 skips, and that the derivation with 5 skips would be inferior in approximately the same ratio. Weighting therefore as above indicated, we obtain values for the median distances as given in the last line of Table XVI. These are the measures that will be used in this study; but they differ so little from those obtained without weighting that the latter may serve almost as well.

Concretely these results mean that if we represent the difference between no spelling ability at all and the ability of typical 3d-grade children by x, the ability of typical 4th-grade children will be represented by $x + 1.351$, of typical 5th-grade children by $x + 1.351 + .836$ or $x + 2.187$, of typical 6th-grade children by $x + 3.238$, of typical 7th-grade children by $x + 3.899$, and of typical 8th-grade children by $x + 4.809$. (The determination of the value of x is not material in the present connection. We shall, however, have something to offer on this point in a later section.) The median distances suggest that so far as spelling is concerned the equal time intervals of one year between the grades do not at all correspond to the differences in ability. The difference between 3d-grade performance and 4th-grade performance is more than twice as great as the difference between 6th- and 7th-grade performance. Whether this is due to a more or less common failure in the 7th grade to give as much attention to spelling as in earlier grades or whether in general 6th and 7th grades are actually closer together than others, is a question which we cannot settle. That a lack of effort to instruct in spelling in the higher grades does not fully account for the differences is suggested by the fact that the 8th grade stands at a greater distance from the 7th than the 5th does from the 4th or the 7th from the 6th.

§ 12. *Scaling the Words*

Assuming that the normal surface of frequency represents the distribution of spelling ability in each grade, we shall now seek to determine how difficult each one of the 50 words listed

in Tables III and IV is for each grade. A word spelled by one hundred per cent of the pupils in, say, the 3d grade would have no difficulty for that grade. The ability of all pupils would be greater than the ability required to spell it, and the entire area of the frequency surface would lie above it—i.e., to the right of it. In Fig. 12, if *OP* represents the Probable Error, it would be located theoretically at an indefinite distance to the left of the point *O*, a distance, however, which we may for practical purposes call 5 or 6 times as great as *OP*—i.e., 5 P.E. or 6 P.E. A word spelled by 98 per cent of the pupils becomes more intelligible. It would be located at a point *K*, a vertical at which (*KL*) would cut off 2 per cent of the area of the entire frequency surface. The point *K* will be found to be at a distance

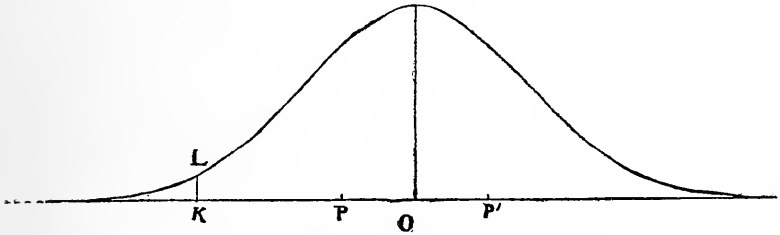


Fig. 12. Normal Surface of Frequency.

of about 3 P.E. below the median *O*, i.e., at 3 P.E. A word spelled by nobody—i.e., a word rated at 0—would be at, say, +6 P.E., and a word spelled correctly by 50 per cent of the group would be located at the median *O*, that is, at a point above and below which are an equal number of cases.

It will be interesting and will serve to show the misleading character of per cent ratings to observe what we mean by saying that one word is more difficult than another. Observe the two following groups of words taken from Tables III and IV for the 3d grade:

(A)	Per Cent Correct	(B)	Per Cent Correct
tailor	38	beautiful	10
lesson	37	beginning	9
another	36	telephone	8
wear	35	pigeons	7

According to the ratings of these words the differences in point of difficulty between the words of group A are equal to the differences in group B, for the differences are all represented by 1 per cent. Habitually we are likely to think that this is true. But such a way of thinking quite neglects the form of distribution of spelling ability. In fact it assumes that the frequency surface is a rectangle—i.e., that there are just as many very poor or very good spellers as there are spellers of medium ability. This we know is not true. The mediocre are always much more numerous than the dull or the gifted. A figure such as Fig. 12 takes account of this fact.

Now the words in group A are much nearer the median (which would be a word 50% correct) than are those of group B. They are located on the base line at points such that between adjacent verticals drawn at these points one per cent of the area will lie. The words of group B, more remotely placed with reference to the median, are also so situated that between their adjacent verticals one per cent of the area will lie. But the points for group B stand at greater distances apart than do the points for group A because the verticals or ordinates are shorter for the remoter group. As a matter of fact, the difference in difficulty between "beautiful" and "pigeons" is more than twice as much as the difference in difficulty between "tailor" and "wear," although each difference is represented by the same per cent.

Bearing in mind the meaning of these per cent values we may readily place the 50 words of Tables III and IV along the x-axis or base line of a normal frequency surface. "Even," which is rated 59 per cent for the 3d grade, would be at a point *below* the median between whose ordinate and the median ordinate is 9% of the area of the surface. Calling the median zero and referring to Table XIV, we find that 9% of the cases (900 in 10,000) corresponds to a value of P.E. which lies between .3 and .35. By interpolation this value is found to be .338. Therefore the position of "even" is at —.338 P.E. This may be represented on Fig. 13 by the point 1. "Lesson" (37% correct) will be at a point *above* zero between which and zero are 13% of the cases of a normal frequency surface. Table XIV locates this point at +.49 P.E. (Point 2, Fig. 13). "Only" (65%

correct) is at —.572 P.E. (Point 3, Fig. 13); "smoke" (46%) at +.148 P.E. (point 4); "pear" (31%) at +.735 P.E. (point 5); "minute" (26%) at +.955 P.E. (point 6); "cousin" (19%) at + 1.300 P.E. (point 7), and so on. Words rated above 50% are located below the median; those under 50% are above the median. Their distances from the median are negative in the first case and positive in the second.

Assuming the same form of distribution for the 4th grade we find that "even" (79%) is located at — 1.20 P.E., "only" (75%) at precisely — 1.00 P.E., and "pear" (42%) at +.30 P.E. Similarly for each grade by using the per cents of Table III and IV and the P.E. equivalents of Table XIV we may "place" all

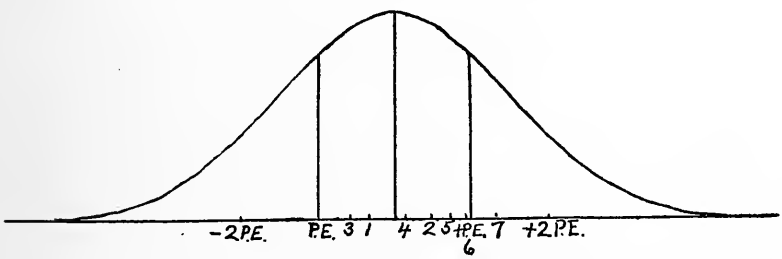


FIG. 13. Showing the placing of the first 7 words of the Preferred List. 3d grade.

the words. Table XVII gives the per cents and P.E. equivalents of the 50 words of the Preferred Lists which from now on will be treated as one list. Figs. 14, 15, 16, 17, 18 and 19 show how the words appear when arranged on a linear scale for each grade. For the meanings of the numbers, each of which refers to a word of the Preferred List, see Table XVII or Appendix II.

Table XVII with its corresponding figures (14 to 19) affords standards for grade performances. As will be observed, the P.E. values of all the words are calculated for each grade with reference to the median of that grade, which is called zero. Their use may be illustrated with reference to the 4th grade. We may test a pupil of that grade by beginning with the easiest word and proceeding to the next hardest and the next and so on. The series would run: 1 *even,* 3 *only,* 2 *lesson,* or 5 *front,* 28 *chicken,* or 41 *Tuesday,* 4 *smoke,* 11 *pretty,* 8 *bought.* . . . By the time

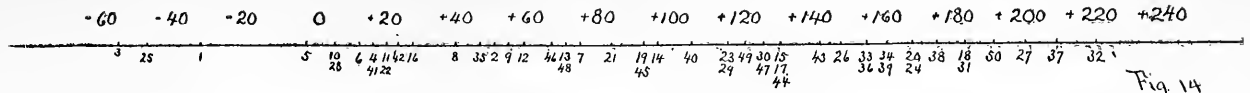


Fig. 14
3rd Grade

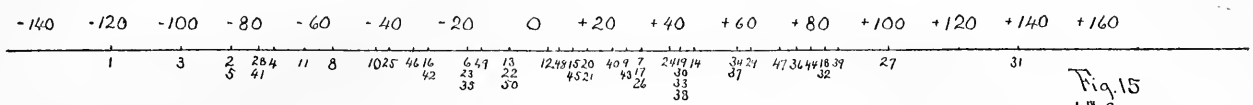


Fig. 15
4th Grade

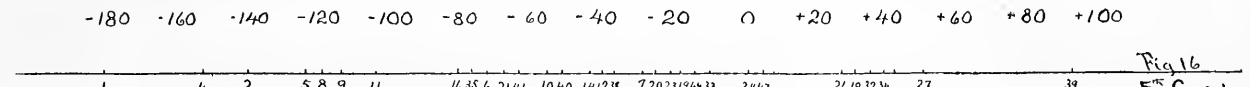


Fig. 16
5th Grade

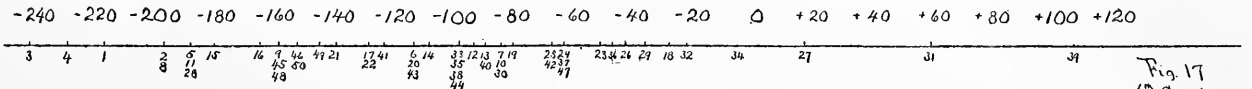


Fig. 17
6th Grade

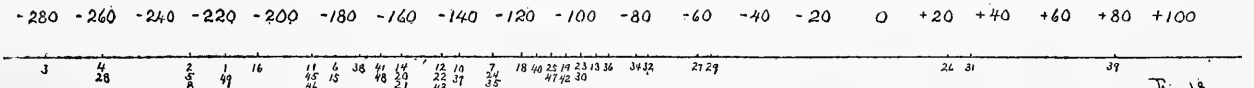


Fig. 18
7th Grade

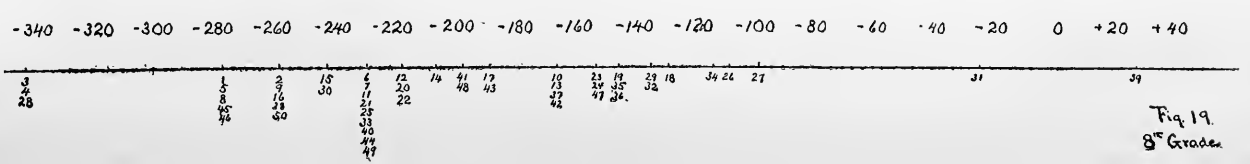


Fig. 19.
8th Grade

TABLE XVII

PER CENTS CORRECT AND P.E. EQUIVALENTS FOR EACH WORD OF THE PREFERRED LIST. GRADES 3–8. SEE FIGS. 14–19

No. of Word	Words	3d Year		4th Year		5th Year		6th Year		7th Year		8th Year	
		%	P.E.	%	P.E.	%	P.E.	%	P.E.	%	P.E.	%	P.E.
1	even	59	— .337	79	—1.196	89	—1.819	93	—2.188	93	—2.188	97	—2.789
2	lesson	37	+ .492	72	— .864	83	—1.415	91	—1.988	94	—2.305	96	—2.597
3	only	65	— .571	75	—1.000	89	—1.819	95	—2.439	97	—2.789	99	—3.450
4	smoke	46	+ .149	69	— .735	85	—1.537	94	—2.305	96	—2.597	99	—3.450
5	front	51	— .037	72	— .864	80	—1.248	90	—1.900	94	—2.305	97	—2.789
6	sure	47	+ .112	55	— .187	69	— .735	78	—1.145	89	—1.819	94	—2.305
7	pear	31	+ .735	42	+ .299	58	— .299	72	— .864	81	—1.302	94	—2.305
8	bought	40	+ .376	65	— .571	79	—1.196	91	—1.988	94	—2.305	97	—2.789
9	another	36	+ .531	43	+ .261	78	—1.145	86	—1.602	94	—2.305	96	—2.597
10	forty	49	+ .037	62	— .453	65	— .571	72	— .864	83	—1.415	87	—1.670
11	pretty	45	+ .187	67	— .652	76	—1.047	90	—1.900	90	—1.900	94	—2.305
12	wear	35	+ .571	49	+ .037	61	— .414	74	— .954	84	—1.475	93	—2.188
13	button	32	+ .693	52	— .074	61	— .414	73	— .909	74	— .954	87	—1.670
14	minute	26	+ .954	38	+ .453	62	— .453	77	—1.096	86	—1.602	92	—2.083
15	cousin	19	+1.302	47	+ .112	69	— .735	89	—1.819	89	—1.819	95	—2.439
16	nails	43	+ .261	58	— .299	71	— .820	87	—1.670	92	—2.083	96	—2.597
17	janitor	19	+1.302	42	+ .299	58	— .299	81	—1.302	81	—1.302	90	—1.900
18	saucer	11	+1.819	29	+ .820	42	+ .299	58	— .299	79	—1.196	81	—1.302
19	stopping	27	+ .909	39	+ .414	55	— .187	71	— .820	76	—1.047	84	—1.475
20	sword	13	+1.670	46	+ .149	57	— .261	78	—1.145	86	—1.602	93	—2.188
21	freeze	29	+ .820	46	+ .149	68	— .693	83	—1.415	86	—1.602	94	—2.305
22	touch	45	+ .187	52	— .074	60	— .376	81	—1.302	84	—1.475	93	—2.188
23	whistle	22	+1.145	55	— .187	56	— .224	64	— .531	75	—1.000	85	—1.537
24	carriage	13	+1.670	40	+ .376	50	— .000	67	— .652	81	—1.302	85	—1.537
25	nor	63	— .492	61	— .414	65	— .571	68	— .693	77	—1.096	94	—2.305
26	already	16	+1.475	42	+ .299	43	+ .261	62	— .453	44	+ .224	77	—1.096
27	beginning	9	+1.988	25	+1.000	37	+ .492	46	+ .149	66	— .612	75	—1.000
28	chicken	49	+ .037	70	— .778	83	—1.415	90	—1.900	96	—2.597	99	—3.450
29	choose	22	+1.145	34	+ .612	48	+ .074	60	— .376	65	— .571	82	—1.357
30	circus	20	+1.248	39	+ .414	50	— .000	72	— .864	75	—1.000	95	—2.439
31	grease	11	+1.819	18	+1.357	37	+ .492	35	+ .571	42	+ .299	57	— .261
32	pigeons	7	+2.188	29	+ .820	41	+ .337	57	— .261	70	— .778	82	—1.357
33	quarrel	15	+1.537	39	+ .414	53	— .112	75	—1.000	86	—1.602	94	—2.305
34	saucy	14	+1.602	35	+ .571	40	+ .376	52	— .074	71	— .820	78	—1.145
35	tailor	38	+ .453	55	— .187	70	— .778	75	—1.000	81	—1.302	84	—1.475
36	telegram	15	+1.537	31	+ .735	39	+ .414	63	— .492	73	— .909	84	—1.475
37	telephone	8	+2.083	35	+ .571	48	+ .074	67	— .652	83	—1.415	87	—1.670
38	tobacco	12	+1.742	39	+ .414	60	— .376	75	—1.000	88	—1.742	96	—2.597
39	too	14	+1.602	28	+ .864	27	+ .909	24	+1.047	30	+ .778	43	+ .261
40	towel	24	+1.047	44	+ .224	64	— .531	73	— .909	78	—1.145	94	—2.305
41	Tuesday	46	+ .149	70	— .778	67	— .652	80	—1.248	87	—1.670	91	—1.988
42	tying	44	+ .224	58	— .299	70	— .778	68	— .693	76	—1.047	87	—1.670
43	whole	17	+1.415	43	+ .261	64	— .531	78	—1.145	84	—1.475	90	—1.900
44	against	19	+1.302	30	+ .778	54	— .149	75	—1.000	84	—1.475	94	—2.305
45	answer	27	+ .909	47	+ .112	67	— .652	86	—1.602	90	—1.900	97	—2.789
46	butcher	33	+ .652	59	— .337	69	— .735	85	—1.537	90	—1.900	97	—2.789
47	guess	20	+1.248	32	+ .693	49	+ .037	67	— .652	77	—1.096	85	—1.537
48	instead	32	+ .693	48	+ .074	62	— .453	86	—1.602	87	—1.670	91	—1.988
49	raise	21	+1.196	54	— .149	67	— .652	84	—1.475	93	—2.188	94	—2.305
50	beautiful	10	+1.900	52	— .074	70	— .778	85	—1.537	94	—2.305	96	—2.597

we have reached 13 *button,* 22 *touch,* 50 *beautiful,* 12 *wear,* and 48 *instead,* we are dealing with a group of words which 50 per cent of 4th-grade children spell correctly. The performance of

a given 4th-grade pupil should approximate at least the standard set by these words. If we are asked, "What is 4th grade spelling ability?" we may answer that it is the ability to spell these words that cluster about the median. Of course it is to be expected that any given pupil will miss some of the easier words and spell some of the harder words. We should test him by the whole series of 50 and his errors for words below the median may be balanced against correct spellings of words above the median at an approximately equal distance. He may miss 49 *raise* (—.15 P.E.) but spell 20 *sword* or 21 *freeze* (+.15 P.E.). He may miss 16 *nails,* but spell 7 *pear.* In such cases he should be credited with having spelled the easier word.

In a similar way, by using the words in the order in which they are placed for any other grade, we may determine whether a child is as good a speller as the median children of that grade. We do not need, however, to use the median as a standard unless we wish to. We may choose + 40 or + 60 and ascertain whether children are able to spell up to that point in the same manner as is indicated above for the zero-point. It is to be observed, however, that our series does not offer a very satisfactory test in the higher grades for such standards, because there are so few words that are placed as high or higher than + 40 or + 60. The words, in short, are not difficult enough for this purpose. In a later section of the paper we shall introduce harder words into the series precisely with the object of affording a fuller test of ability for the higher grades.

There remains, however, for the present one other use which may be made of our data. We may wish to disregard grades altogether and seek an answer to the question, In general, how hard are these words for children of the elementary school above the 2d year? or, with reference to a graphical representation, What is the average position of each word on a linear scale—that position from which the positions for each grade deviate by the smallest amounts?

To answer such a question we shall have to use one point of reference for all grades instead of a different one for each grade. In the above treatment we have expressed each word-value as a deviation from the median of the particular grade we were considering. We shall now use this same data but

transfer the point of reference to the third-grade median by using the median intervals which were derived in Section 11. In Table XVI (page 39) we have given the results of our inquiry into the amounts of these intervals as follows:

From M_3 to M_4	1.351	P.E.
" M_3 " M_5	2.187	"
" M_3 " M_6	3.238	"
" M_3 " M_7	3.899	"
" M_3 " M_8	4.809	"

Table XVIII gives, for each of the 50 words, its position for each grade when referred to the 3d-grade median as the zero-point or point of reference, together with the "average position" of each word. The method of securing these figures may be illustrated as follows:

From Table XVII the P.E. values of the word "even" for each grade, referred to its own median, are shown to be

3d grade,	— .337
4th "	—1.196
5th "	—1.819
6th "	—2.188
7th "	—2.188
8th "	—2.789

The first of these values is of course already referred to the 3d-grade median. To refer the others to the same point we must increase each of them by the amount by which each grade median stands above the 3d-grade median, i.e., we must find the sum (algebraic) of —1.196 and 1.351, of —1.819 and 2.187, of —2.188 and 3.238, of —2.188 and 3.899, and of —2.789 and 4.809. These sums give the figures of Table XVIII for the word "even." Their arithmetical mean is taken as the average position.

Fig. 20 shows the averages of Table XVIII when reduced to a scale. The noticeable thing about these tabular and graphic representations is the fact that the words from easiest to hardest differ so little. The words "even" and "only" (No. 1 and No. 3), which are —.337 and —.571 respectively for the 3d grade, appear above the zero point at +.699 and +.569. Similarly the word "too" (No. 39) which for the 8th grade alone is + 5.07 becomes for all grades only + 3.491. It is a fact

TABLE XVIII

THE POSITION OF EACH WORD IN EACH GRADE WHEN REFERRED TO THE 3D-GRADE MEDIAN AS THE ZERO-POINT; AND THE AVERAGE POSITION OF EACH WORD FOR ALL GRADES, WHEN SO REFERRED. 1=P.E.

Word Number	Word	3d Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade	Average Position
1	even	—.337	.155	.368	1.050	1.711	2.020	.699
2	lesson	.492	.487	.772	1.250	1.594	2.212	1.135
3	only	—.571	.351	.368	.797	1.110	1.359	.569
4	smoke	.149	.616	.650	.933	1.302	1.359	.835
5	front	—.037	.487	.937	1.338	1.594	2.020	1.057
6	sure	.112	1.164	1.452	2.093	2.080	2.504	1.568
7	pear	.735	1.650	1.888	2.374	2.597	2.504	1.958
8	bought	.376	.780	.991	1.250	1.594	2.020	1.169
9	another	.531	1.612	1.042	1.636	1.594	2.212	1.078
10	forty	.037	.898	1.616	2.374	2.484	3.139	1.758
11	pretty	.187	.699	1.140	1.338	1.999	2.504	1.311
12	wear	.571	1.388	1.773	2.284	2.424	2.621	1.844
13	button	.693	1.277	1.733	2.329	2.945	3.139	2.026
14	minute	.954	1.804	1.734	2.142	2.297	2.726	1.943
15	cousin	1.302	1.463	1.452	1.419	2.080	2.370	1.681
16	nails	.261	1.052	1.367	1.568	1.816	2.212	1.379
17	janitor	1.302	1.650	1.888	1.936	2.597	2.909	2.047
18	saucer	1.819	2.171	2.486	2.939	2.703	3.507	2.604
19	stopping	.909	1.765	2.000	2.418	2.852	3.334	2.213
20	sword	1.670	2.500	1.926	2.093	2.297	2.621	2.185
21	freeze	.820	1.500	1.494	1.823	2.297	2.504	1.740
22	touch	.187	1.277	1.811	1.936	2.424	2.621	1.709
23	whistle	1.145	1.164	1.963	2.707	2.899	3.272	2.193
24	carriage	1.670	1.727	2.187	2.586	2.597	3.272	2.340
25	nor	—.492	.937	1.616	2.545	2.803	2.504	1.652
26	already	1.475	1.650	2.448	2.785	4.123	3.713	2.699
27	beginning	1.988	2.351	2.679	3.387	3.287	3.809	2.917
28	chicken	.037	.573	.772	1.338	1.302	1.359	.897
29	choose	1.145	1.963	2.261	2.862	3.328	3.452	2.502
30	circus	1.248	1.765	2.187	2.374	2.899	2.370	2.141
31	grease	1.819	2.708	2.679	3.809	4.198	4.548	3.294
32	pigeons	2.188	2.171	2.524	2.977	3.121	3.452	2.739
33	quarrel	1.537	1.765	2.075	2.238	2.297	2.504	2.069
34	saucy	1.602	1.922	2.563	3.164	3.079	3.664	2.666
35	tailor	.453	1.164	1.409	2.238	2.597	3.334	1.866
36	telegram	1.537	2.086	2.601	2.746	2.990	3.334	2.549
37	telephone	2.083	1.922	2.261	2.586	2.485	3.139	2.413
38	tobacco	1.742	1.765	1.811	2.238	2.157	2.212	1.988
39	too	1.602	2.215	3.096	4.285	4.677	5.070	3.491
40	towel	1.047	1.575	1.656	2.329	2.754	2.504	1.978
41	Tuesday	.149	.573	1.535	1.990	2.229	2.821	1.550
42	tying	.224	1.052	1.409	2.545	2.852	3.139	1.870
43	whole	1.415	1.612	1.656	2.093	2.424	2.909	2.018
44	against	1.302	2.129	2.038	2.238	2.424	2.504	2.106
45	answer	.909	1.463	1.535	1.636	1.999	2.020	1.594
46	butcher	.652	1.014	1.452	1.701	1.999	2.020	1.473
47	guess	1.248	2.044	2.224	2.586	2.803	3.272	2.363
48	instead	.693	1.425	1.734	1.636	2.229	2.821	1.756
49	raise	1.196	1.202	1.535	1.763	1.711	2.504	1.652
50	beautiful	1.900	1.277	1.409	1.701	1.594	2.212	1.682

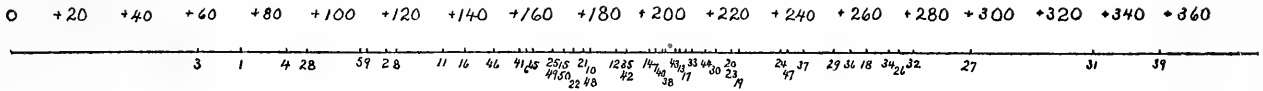


FIG. 20. The positions of each word of the Preferred List on a scale of difficulty for grades 3-8 together, 0 being the 3d-grade median and 100 being P.E. The Preferred List is as follows:

1. even	11. pretty	21. freeze	31. grease	41. Tuesday
2. lesson	12. wear	22. touch	32. pigeons	42. tying
3. only	13. button	23. whistle	33. quarrel	43. whole
4. smoke	14. minute	24. carriage	34. saucy	44. against
5. front	15. cousin	25. nor	35. tailor	45. answer
6. sure	16. nails	26. already	36. telegram	46. butcher
7. pear	17. janitor	27. beginning	37. telephone	47. guess
8. bought	18. saucer	28. chicken	38. tobacco	48. instead
9. another	19. stopping	29. choose	39. too	49. raise
10. forty	20. sword	30. circus	40. towel	50. beautiful

then that *for these words* the influence of higher grades is to make easy words harder and of lower grades to make hard words easier. That is, grade considered, these words are harder for children of the upper grades than they are for those of the lower grades. There are at least two reasons for this condition.

First, as a rule these particular words are taught in the lower grades. A popular speller, taken at random, presents 31 of the 50 words in the 3d year's work, 10 in the 4th, 2 in the 5th, and none in higher grades. There is nothing to lead one to suppose that this is peculiar. The words were among those chosen, it will be remembered, as at least in the speaking vocabulary of 3d-grade children. Most of them if taught at all will be taught in that grade. We may assume therefore that the 3d-grade record is somewhat affected by the recency with which these words have been presented. The succeeding grades will to some extent be discriminated against in the record.

Second, the necessary basis of selection for these words from the larger lists would make it impossible for the words to take the same position on the scale for all grades. Consider a word which was spelled correctly by 50% of the 3d-grade children. Such a word would be at M_3. In order to take the same position on the 8th-grade record it must be as far below M_8 as is the distance, already determined, between M_3 and M_8 or -4.809 P.E. To do this it would have to be spelled correctly by 9994 pupils out of 10,000 (Table XIV), i.e., it would be 100% correct. But such a word would not have been selected, because it is not difficult enough in the 8th grade to be of any value as a test of ability. On the other hand, a word missed often enough in the 8th grade to be satisfactory as a test (say, 90% correct) would have to be less than 1% correct on the 3d-grade record in order to take the same position on the scale. Such a word would have been of no use as testing 3d-grade ability and would have been rejected.

The fact is that the span from 3d to 8th grade is—if our median distances be correct—too great for any list of words to be in all respects satisfactory. We need several lists each of which shall be given to three or four consecutive grades and overlapping on one another—e.g., one for 2d, 3d, and 4th grades, another

for 3d, 4th, 5th and 6th grades, and another for 5th, 6th, 7th and 8th grades. An attempt will be made in a later section to do this and to show the results that may be expected.

§ 13. *The Use of the Scale*

Meanwhile, however, we venture to think that the scale as shown in Fig. 20 is important and valid within its range. It may be used in several ways of which at least three are important.

1. It may be used just as it is without reference to the fact that the words are not separated from each other by equal intervals. We know the value or weight to assign to each word. We shall therefore not make the mistake of assuming that all the words are of the same value, as is the usual school practice.

2. Certain words of the series may be used which differ from each other by approximately equal steps.

TABLE XIX

WORDS ARRANGED IN ORDER OF DIFFICULTY ACCORDING TO THE SCALE AND THEIR P.E. VALUES

No. on Scale	Word	P.E. x 100	No. on Scale	Word	P.E. x 100
3	only	57	14	minute	194
1	even	70	7	pear	196
4	smoke	84	40	towel	198
28	chicken	90	38	tobacco	199
5	front	106	43	whole	202
9	another	108	13	button	203
2	lesson	114	17	janitor	205
8	bought	117	33	quarrel	207
11	pretty	131	44	against	211
16	nails	138	30	circus	214
46	butcher	147	20	sword	219
41	Tuesday	155	23	whistle	219
6	sure	157	19	stopping	221
45	answer	159	24	carriage	234
25	nor	165	47	guess	236
49	raise	165	37	telephone	241
15	cousin	168	29	choose	250
50	beautiful	168	36	telegram	255
22	touch	171	18	saucer	260
21	freeze	174	34	saucy	267
10	forty	176	26	already	270
48	instead	176	32	pigeons	274
12	wear	184	27	beginning	292
35	tailor	187	31	grease	329
42	tying	187	39	too	349

3. Small groups of words may be so selected as to be equally difficult as groups; or they may be so selected that their group-difficulties constitute an ascending series from easy to hard, differing by equal amounts.

1. By the first of these methods the entire series would be utilized or so much of it as in any given case would thoroughly test the subject. The order of the words of the series as given in Figure 20 is shown in Table XIX in the first column and in the second column the test values or weights of these words are given.

2. If it is desired to use a scale whose words differ in difficulty by equal steps, the arrangement as shown in Table XX will be found convenient.

TABLE XX
A Ten-point Scale

No. of Word (Fig. 20)	Word	P.E. x 100	Δ
3	only	57	
			27
4	smoke	84	
			24
9	another	108	
			23
11	pretty	131	
			28
45	answer	159	
			28
35	tailor	187	
			27
30	circus	214	
			27
37	telephone	241	
			26
34	saucy	267	
			25
27	beginning	292	

To this series may be added 39 "too" whose P.E. x 100 is 349 and which differs from "beginning" by 57 or approximately two steps.

In the series of Table XX the average step is 26.2 with an A.D. of 1.3; or if the word "too" is included the average step is 26.6 with an A.D. of 1.4. This is quite accurate enough for any use to which the scale is likely to be put. If this conclusion

is accepted, these eleven words may be used to express our judgments of other words concretely and in terms that everybody can understand. We should not then have to resort to such terms as "hard," "easy," "rather difficult," "very hard," etc., but we may judge a word to be "as hard as 'another,'" "equal in difficulty to 'beginning,'" "as hard as 'answer' but not as hard as 'tailor,'" etc. It is very desirable that other words should at some time be added to the scale at both ends. There are many words harder than "beginning" or "too" and there are others easier than "only," although the latter do not constitute much of a school problem. Neither set, however, could be used over a range as wide as 3d to 8th grades.

3 (a). It is often desirable to offer tests of equal difficulty, but of different words at various intervals of time to the same group or to the same individual. We may thus secure a progress record. In spelling, however, this has proved to be very difficult if not impossible. We can never be sure that the second or third test is equal in difficulty to the first test. In fact we may be pretty sure it is not. To give the same words over again is often valueless because of the added special familiarity with them. The following lists therefore are offered as lists of equal difficulty. The sum of the P.E. values in each is 976 or 977. In using them the words may be weighted as indicated, or may, with no great loss in precision, be each given a credit of 1.

Number in Preferred List		Weight	Number in Preferred List		Weight
	Group A			Group B	
41	Tuesday	16	45	answer	16
10	forty	18	48	instead	18
40	towel	20	43	whole	21
44	against	22	17	janitor	21
47	guess	24	24	carriage	24
	Group C			Group D	
49	raise	17	21	freeze	18
22	touch	17	12	wear	19
42	tying	19	7	pear	20
14	minute	20	13	button	21
18	saucer	27	20	sword	22
	Group E			Group F	
16	nails	14	8	bought	12
46	butcher	15	11	pretty	13
15	cousin	17	19	stopping	23
29	choose	26	37	telephone	25
32	pigeons	28	34	saucy	27

(b) It may also be desirable to test, not with single words, which only in the long run may be expected to conform to the positions assigned to them, but with groups of words whose difficulties as groups differ by constant amounts. Such a series of groups arranged from easy to hard would themselves constitute a scale—a sort of Binet-Simon scale for measuring ability in spelling. On the analogy of the Binet-Simon scale we might easily fix upon a certain minimum performance for a group at which or better than which a subject might be allowed to have "cleared" that group and might pass on to the next. He might also be given additional credits for spelling words in groups above the highest one which he cleared.

The groups are arranged in order of difficulty, Group I being the easiest and Group VII the hardest. Within each group the four words are also arranged in their order of difficulty beginning with the easiest. Since, however, within each group the words differ little in difficulty, they may be taken as having equal weights without material error. It is true that Group VII is not nearly as satisfactory in this respect as the others, differing between the first and fourth words by 1.08 P.E., whereas the first six groups have a range of but .225 on the average.

	Group I	P.E. x 100
3	only	57
1	even	70
4	smoke	84
28	chicken	90
	Average	75

	Group II	P.E. x 100
5	front	106
9	another	108
2	lesson	114
8	bought	117
	Average	111

	Group III	P.E. x 100
16	nails	138
46	butcher	147
41	Tuesday	155
6	sure	157
	Average	149

	Group IV	P.E. x 100
10	forty	176
12	wear	184
42	tying	187
38	tobacco	199
	Average	186.5

	Group V	P.E. x 100
33	quarrel	207
30	circus	214
24	carriage	234
47	guess	236
	Average	223

	Group VI	P.E. x 100
29	choose	250
36	telegram	255
34	saucy	267
26	already	270
	Average	260.5

	Group VII	P.E. x 100
37	telephone	241
32	pigeons	274
31	grease	329
39	too	349
	Average	298

The average differences in difficulty between these groups in succession are 36, 38, 37.5, 36.5, 37.5 and 37.5. This is probably the most important use of the scale, for present school practice.

If it is true that the general scale (Table XVIII and Fig. 20) may be used in these three ways—as a whole, by words selected to be at equal intervals, and by grouping words so that the groups are equal or differ by equal amounts—then it is also true that each of the grade scales (Figs. 14-19) may be used in like manner each for the grade to which it applies. It is probably true, moreover, that the grade scales will more closely fit real conditions in any given instance than will the scale for all grades. The labor of making selections and groupings of words for these scales is not great and may be made by any one on the analogy of the method used above.

§ 14. *The Zero-Point of Spelling Ability*

As has been suggested in previous sections, we have only succeeded in scaling by means of these 50 words a limited segment of the entire projection representing spelling ability. Our list is essentially an easy list, testing that ability only to a moderate degree. Words like "fatiguing," "guarantee," and "conscientious" (Rice Sentence Test) would stand much higher in the scale and require a considerable extension of it to the right; while such unfamiliar words as "eurycerous," "delitescence," and "gallinaceous" (Klein, '12, pp, 388, 389) would take still higher positions, passing quite beyond the range of the ability of elementary-school children.

On the other hand, our scale is as certainly limited at the low end. There are many easier words than any we have used so far. Such words would reach far down on the scale towards the place where the absolute zero-point lies. But they would have been totally unfit for use in the higher grades. In fact, with the wide range of ability between 3d and 8th grades, it is surprising that we find any words at all which will afford a test at both extremes.

Without seeking to determine the limit of the high end of the scale—perfect spelling ability—it is quite possible, and theoretically very desirable, to find the limit of the low end, i.e.,

to find the point where spelling ability just begins to be a positive quantity.

How far, then, below the 3d-grade median, which has hitherto been our point of reference, is the absolute zero-point?

In order to answer this question, a test was given to children of the 2d, 3d, and 4th grades. It consisted of 50 words in sentences. Nineteen of these had already been used in the Selected List (100 word list); and, of these, 6 had been chosen for the Preferred List. They had all been spelled, about 40 per cent or more correct, by the third-grade children. The remaining 32 words were thought to be among the easiest in the language: he, is, on, the, to, of, for, day, etc.

They were put into sentences as follows and dictated at schools II and VIII:

EASY 50-WORD TEST

1. *You will hear him coming.*
2. *He is on the road and* is *almost sure to pass in front of me.*
3. *I send for* him *every day.*
4. *Go into* the *school.*
5. *But do not touch* the *table.*
6. He *also has only one pair* of *shoes.*
7. *They are* not *at all pretty.*
8. *No man ought* to *steal even a penny.*

It seems clear that a child who cannot spell *any one* of these words has substantially no spelling ability. Since our study is limited to written words we shall say, therefore, that for our purpose a child who does not, save by chance, write a single word of this list so that it can be recognized as correctly spelled has no ability.

On account of the marked improvement in spelling of children in the latter half of the second school year over those in the first half of that year, we have treated the two half-years of the 2d grade separately, calling the lower half *2a* and the upper *2b*. We shall proceed as follows. We shall find the distance between the 3d-grade median and the *2b*-grade median and the distance between the latter and the *2a*-grade median. Then if there are children of the *2a* grade who utterly break down and

fail to write any word correctly we shall find their place in the *2a* distribution.

Table XXI shows the records of individual pupils according to their rating in the Easy 50-Word Test. Note the fact that no children of the 4th, 3d, or *2b* grades wholly failed in this test. In *2a,* however, 39 children were rated 10% or less, and of these there were 8 pupils who were actually marked zero. This is 4.6% of all the children of *2a.*

TABLE XXI

DISTRIBUTION OF INDIVIDUAL RATINGS. EASY 50-WORD TEST

Table reads: in *2a* 39 children, or 22%, were rated between 0 and 10%; 32 children, or 18%, were rated between 11% and 20%, etc. In *2b* 5 children, or 3%, were rated between 11% and 20%, etc.

Per Cent Correct	*2a* Grade		*2b* Grade		3d Grade		4th Grade	
	No.	%	No.	%	No.	%	No.	%
0– 10	39	22	0	0	0	0	0	0
11– 20	32	18	5	3	1	.6	0	0
21– 30	37	21	9	5	4	2	0	0
31– 40	27	16	29	17	7	4	0	0
41– 50	18	10	26	15	11	7	4	1
51– 60	14	18	47	28	25	15	13	4
61– 70	5	3	31	18	33	20	29	9
71– 80	2	1	14	8	36	21	50	16
81– 90	1	.6	7	4	30	18	86	27
91–100	0	0	1	.6	21	13	134	42
Totals	175		169		168		316	
Medians		26.50		56.17		72.50		88.12

The medians for the grades are as follows: for *2a,* 26.50%; for *2b,* 56.17%; for 3d grade, 72.50%; and for 4th grade, 88.12%. The rapid rise of spelling ability from low second through the fourth grade is remarkable. It is much greater than the improvement during the next four years. Although the interval in time between *2a* and *2b* is but half a year, the medians suggest that the increase in ability between these grades is greater than it is between any consecutive yearly grades above the second. Further analysis will more precisely confirm this inference.

Proceeding as in the case of grades 3 to 8, we show in Table XXII the amount and per cent of overlapping of each

grade beyond the medians of the other grades, together with the corresponding linear segment in terms of the Probable Error as a unit.

TABLE XXII

AMOUNT AND PER CENT OF OVERLAPPING WITH P.E. EQUIVALENTS. EASY 50-WORD TEST

		2*a* Grade	2*b* Grade	3d Grade	4th Grade
2*a* grade.....	No.	....	17	3	0
	%	....	9.71	1.71	0
	P.E.	....	1.9254	3.1429	?
2*b* grade.....	No.	160	....	22	4
	%	94.67	....	13.02	2.37
	P.E.	2.3932	....	1.6690	2.9395
3d grade.....	No.	165	140	....	30
	%	98.21	83.33	....	17.86
	P.E.	3.1143	1.4341	....	1.3649
4th grade.....	No.	316	308	262	....
	%	100	97.47	82.91	....
	P.E.	?	2.8976	1.4094	....

From these results, Table XXIII is computed. The object in this table is to show how, by using all the data of Table XXII, various values of the median intervals may be obtained whose averages will be the most probably correct values. The interval between the medians of *2a* and *2b* is written *M2a-2b;* that between the medians of *2b* and the 3d grade is written *M2b*-3; etc.

It may be remarked parenthetically that in the number 1.3771 of Table XXIII for the difference between the 3d- and 4th-grade medians, we have a striking confirmation of the substantial accuracy of our results as shown in Table XVI. The corresponding number is there given as 1.3505. That these should differ by so little when carried out upon different test material is exceedingly satisfactory.

According to Table XXIII, the *2b*-grade median is approximately 1.35 P.E. below the 3d-grade median. The *2a*-grade median is about 1.87 P.E. further below, or 3.22 P.E. below the 3d-grade median which we have thus far used as our origin or point of reference.

But we have not yet reached the point of zero ability. Typical

TABLE XXIII

VALUES OF MEDIAN INTERVALS AND THEIR DERIVATION
(2*a*–4TH GRADE)

	M_{2a-2b}	M_{2b-3}	M_{3-4}
	1.9254 (direct)	1.6690 (direct)	1.3649 (direct)
	2.3932 (direct)	1.4341 (direct)	1.4094 (direct)
	1.4739 (M_{2a-3}—M_{2b-3})	1.2175 (M_{2a-3}—M_{2a-2b})	1.2705 (M_{2b-4}—M_{2b-3})
	1.6802 (M_{3a-2}—M_{3-2b})	1.5746 (M_{2b-4}—M_{3-4})	1.4635 (M_{4-2b}—M_{3-2b})
		.7211 (M_{3-2a}—M_{2b-2a})	
		1.4882 (M_{4-2b}—M_{4-3})	
Averages	1.8682	1.3518	1.3771

2a children have *some* ability, namely, according to our record, an ability to score 26.5% in the Easy 50-Word Test. The children of that grade who were unable to write any word correctly were 8 in number, representing 4.6 per cent. These 8 are included in the 39 rated between zero and 10% (Table XXI). Assuming that *2a* children are grouped about their median according to the "normal" distribution, we find that in order to cut off 4.6% from the low end we must take a point 2.5 P.E. below the median, (See Table XIV). This brings the zero-point at 5.72 P.E. below the 3d-grade median (3.22 + 2.5).

This figure, 5.72 P.E., can only be taken as approximately correct. It would be somewhat influenced by an increase of the number of children tested. There is, however, no reason to suppose that the children in schools II and VIII were unusual. The testing in grades 3 to 8 in all other schools shows that results in these two schools do not materially differ from the general results. In both central tendencies and variabilities they are a fair average. There seems, then, to be no good reason why we should not consider the ratings of children in these

schools as typical. It must be borne in mind, however, that the classification of children into grades is a broad one. Just as we found it necessary to treat 2d-year children in half-yearly sections, so we should find that testing at the beginning even of a 20-week term would yield results, especially in the low grades, quite different from those obtained by testing toward the close of the term. Accordingly, the middle of the term is the best time at which to find typical conditions. Moreover, in order that the results may be comparable, the testing of all grades should be done at the same time. If *2a* children were tested towards the end of their term in that grade, while *2b* children were tested towards the beginning of theirs, the median interval would be unduly shortened. A considerable addition to the reliability of our results is found in the fact that all children were tested during the 10th week of a 20-week term.

We may therefore conclude that the intervals between grades *2a, 2b,* 3 and 4 are substantially as found in Table XXIII. But as to the distance of the zero-point below the *2a*-grade median, we cannot be precise. Four and six-tenths per cent of the *2a* children got no word right. As many as 22 per cent wrote less than 6 words correctly. Some of them probably spelled these few simple words correctly by mere chance. If this were true, they would have practically no spelling ability. The situation is more complicated than the above analysis indicates. If we were to assume that *all* the children who wrote 0-5 words correctly had practically no spelling ability (written), the zero-point would then be but 1.15 P.E. below the median instead of 2.5 P.E. If we were to assume that *some* of these children—say those who wrote no more than 3 words correctly—had zero ability, we should find that 29 of the 39 in Table XXI were included. Such an assumption would place our zero-point at 1.44 below the median. There are reasons for thinking that this is not far from the true position. The best judgment, therefore, that we can make from our data and from our knowledge and experience of school conditions is, that the zero-point is about 1.5 P.E. below the *2a* median, or about 4.72 P.E. below the 3d-grade median.

We may summarize our results, then, in Table XXIV and Fig. 21, as follows:

TABLE XXIV

MEDIAN INTERVALS. ZERO TO 8TH GRADE MEDIAN

	Successive Intervals	Distance Above 0
2*a* grade	1.50	1.50
2*b* "	1.87	3.37
3d "	1.35	4.72
4th "	1.35	6.07
5th "	.84	6.91
6th "	1.05	7.96
7th "	.66	8.62
8th "	.91	9.53

Fig. 21, page 62, shows these facts graphically.

§ 15. *Observations on the Distributions Shown in Fig. 21*

It is to be remembered that in Fig. 21 the eight surfaces of frequency constructed on each median vertical are theoretical and not according to the record. Moreover, they express the assumptions that for each grade the distribution of ability in spelling is strictly "normal" and that the real variability is alike in all grades. In a later section we shall take up the matter of applying to our results distributions which are not normal. Meanwhile, however, it will be interesting to observe how satisfactory a strictly normal form of distribution proves to be. To the extent that it expresses the same or nearly the same facts as the record (so far as it should, if valid, do so), it shows its value.

1. In Fig. 21 the 2*a* surface of frequency does not reach the 4th-grade median; but it only falls short a little. According to the record in the Easy 50-Word Test no 2*a* child did as well as the median 4th grade child. But the best 2*a* record was 82% which is only a little less than M_4 (88.12 by Table XXI).

2. In the graphic showing the 3d-grade distribution does not quite reach the 8th-grade median. Similarly the record shows that no 3d-grade child obtained a score equal to 94.68 which (Table XI, p. 27) is M_8 for the Selected 100-Word List; although 3 third-grade pupils had scores in the 91 to 95 group. (Table XI, column 2.)

3. By the figure we see that the low end of the 8th-grade distribution falls short of M_3 but not of M_4. In the record the

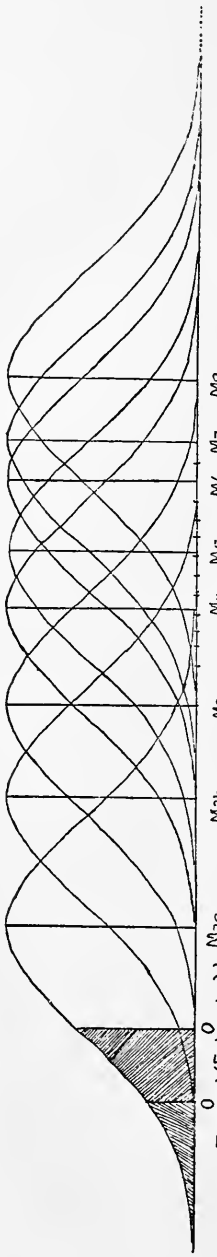


FIG. 21. Distributions (Normal) around grade medians. Entire range of spelling ability in the elementary school. The numbers 3, 4, 9, 11, 45, etc., refer to the words of the ten-point scale; they are placed so as to show the relation of this ten-point scale to the entire scale of difficulty of spelling words. 3 = only, 4 = smoke, 9 = another, 11 = pretty, 45 = answer, 35 = tailor, 30 = circus, 37 = telephone, 34 = saucy, 27 = beginning, and 39 = too.

same is true, although the poorest 8th-graders surpassed the median of the 3d grade by more than the figure would indicate (See Table XI, page 27).

4. The *2b* curve at its low end does not reach zero (record). No 3d-grade child was rated 0 in the Easy 50-Word Test (Table XXI, p. 57).

5. The most remarkable thing in the figure is the fact that the high end of the *2a*-grade distribution extends above the low end of the 8th-grade distribution. Even with all the recent information about retardation, acceleration, mental defect, and precocity combined with mis-grading and forced promotions, some critics will be hardly prepared to believe that any children classified in the first half of the second year of school life can do as well in spelling as the poorest who are classified in the 8th (last elementary) year. This, however, the figure shows; and there is good evidence in our record to support it. We cannot compare these extreme grades directly because they did not write the same test. But in Table XI (page 27) we observe that for the Selected List one 8th-grade pupil is in the 56-60 group. As a matter of fact his paper was rated at exactly 56%. Therefore all the 3d-grade children in groups above 51 to 55 did as well or better. This proves to be 22.5% of the 3d-grade children. It is remarkable that between one-quarter and one-fifth of our 3d-grade children do as well or better than the poorest 8th-grade child. But this is not all. In the Easy 50-Word Test the highest *2a* score was 82%. Only 38 out of 168 3d-grade children, or 22.6%, did better. See Table XXI, page 57. (The one child in group 81-90 in the *2a* column was rated 82. In the 3d-grade column, of the 30 in group 81-90 and the 21 in group 91-100 combined, 38 were rated above 82%.) We therefore find that the best *2a* child and the poorest 8th-grade child are equalled or surpassed by the *same group* of 3d-grade children,—i.e., 22½%. Hence the best *2a* child and the poorest 8th-grader of our record *do* have the same ability. To show this fact, the *2a* curve should pass slightly beyond the low end of the 8th-grade curve, as in fact it does. This may seem to be drawing over-fine conclusions, but it is probable that the real overlapping is as great as the record shows. If, out of so few *2a* children (175), one was found who scored 82%,

it is likely that, with a much larger number, there would not only be more children at 82% but some at even higher rating. Similarly it is likely that if a much greater number of 8th-year children had been tested some would have obtained less than 56%. To the extent that either one or both of these probabilities were true, the overlapping would somewhat exceed that which the record suggests.

6. After the argument of the last paragraph it need hardly be said that there is both in the graphic representation by theory and in the actual record an overlapping of every grade distribution on every other from low 2d to 8th.

The location of the zero-point enables us to draw some interesting conclusions which were not possible before. A few of these will be briefly stated. Taking the medians of each grade as indicating typical abilities, regarding the *estimated* zero-point as the true one, assuming normal distributions and equal real variabilities for all grades, and defining " to spell twice as well " as *"to spell words of twice as much difficulty,"* Fig. 21 shows that *2b* children are more than twice as good spellers as *2a* children, and that 3d-graders are about three times as capable, and 4th-graders 4 times as capable. Fifth-grade children spell twice as well as *2b* children. Eighth-grade children are only twice as good as 3d-grade children. This last statement means that typical children who have reached the 3d grade have half as much spelling ability as is required of the average child in the last year of the elementary school.

We have in Fig. 21 a representation by which the entire range of difficulty of spelling words, appropriate to the elementary school, may be shown. The notation below the base-line shows the positions within that range taken by the 10 words of our scale which stand at equal intervals upon it (Table XX, page 52). Since these words include both the easiest word (" only ") and the hardest (" too ") of the entire scale they show its total spread. It will therefore be seen to what extent the statement is true which was made at the beginning of Section 14 to the effect that with these 50 words we have only succeeded in scaling " a limited segment of the entire projection representing spelling ability." By actual measurement it appears that this segment is but a trifle more than one-fifth of the entire

projection. It will now, however, become apparent that no greater segment *could have been so scaled* reliably from the limited material at our disposal. By reference to Fig. 21 it will be seen that no words easier than "only" could have been used and still have been clearly within the 8th-grade distribution. On the other hand no word that scales much higher than "too" could have been used and still have been clearly within the 3d-grade distribution. With, say, ten thousand children of each grade tested each with a longer list, a wider spread could have been obtained. The series of fifty words which we used spreads over nearly the whole of the base-line common to both the 3d- and 8th-grade curves.

It is evident, therefore, that much remains to be done to perfect a scale which shall pretend to completeness. Some of this further scaling will be undertaken in a later section of this monograph. A great deal must be left for later studies. A great many more words must be used both to fill in the gaps within the present scale and to extend its limits. Our main purpose has been to show the theory and technique required.

§ 16. *Supplementary Testing at Schools VI and VII*

After the data thus far given were in hand the same test material was used in two other schools. The results of this supplementary testing are now given. The Selected List (100 words) was dictated during the fall term of 1912 at schools VI and VII to 1770 children. Two of the assistants in psychology at Teachers College acted as examiners, and the papers were then scored for individual ratings, but not for word ratings. Table XXV gives the distribution of these ratings and the grade medians for these two schools. Table XXVI gives the comparisons by grades of the combined results in schools VI and VII with those in schools II, III, IV, and V taken together. The comparison shows that in general schools VI and VII did not do so well as those tested earlier. On the average the grade medians are nearly 6½ per cent lower.

To one who would expect close conformity to our previous individual ratings in the case of any school taken at random, this discrepancy will be disappointing. But to one who recognizes the wide variability among schools in every subject, the

TABLE XXV

DISTRIBUTION OF INDIVIDUAL RATINGS, SCHOOLS VI AND VII. SELECTED (100) LIST

Per Cent Correct	3d Grade		4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
0– 10	69	22.7	10	3.4	1	.3	3	1.2	0	0	0	0
11– 20	58	19.1	32	10.8	5	1.7	5	2.0	2	.6	0	0
21– 30	65	21.4	29	9.8	12	4.1	8	3.2	2	.6	0	0
31– 40	49	16.1	34	11.5	24	8.3	7	2.8	5	1.5	2	.7
41– 50	29	9.5	40	13.5	31	10.6	13	5.2	3	.9	1	.3
51– 60	17	5.6	51	17.6	41	14.1	23	9.2	13	3.9	3	1.0
61– 70	10	3.3	39	13.2	47	16.2	28	11.2	41	12.2	15	5.1
71– 80	6	2.0	29	9.8	53	18.3	51	20.4	56	16.6	16	5.5
81– 90	1	.3	25	8.4	58	20.0	51	20.4	114	33.8	97	33.1
91–100	0	0	7	2.4	18	6.2	61	24.4	101	30.0	159	54.3
Totals...	304		296		290		250		337		293	
Medians.		25.27		51.75		67.67		79.36		84.86		91.96

TABLE XXVI

COMPARISON OF RESULTS OBTAINED IN SCHOOLS VI AND VII WITH THOSE OBTAINED IN SCHOOLS II, III, IV AND V

Figures show per cent of pupils in each grade who were rated 0–10, 11–20, etc.

Per Cent Correct	3d Grade		4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	Sch. II, III, IV, V	Sch. VI, VII	Sch. II, III, IV, V	Sch. VI, VII	Sch. II, III, IV, V	Sch. VI, VII	Sch. II, III, IV, V	Sch. VI, VII	Sch. II, III, IV, V	Sch. VI, VII	Sch. II, III, IV, V	Sch. VI, VII
0– 10.....	6.9	22.7	.4	3.4	.6	.3	0	1.2	0	0	0	0
11– 20.....	15.2	19.1	4.7	10.8	.6	1.7	0	2.0	0	.6	0	0
21– 30.....	20.5	21.4	7.7	9.8	3.5	4.1	.5	3.2	0	.6	0	0
31– 40.....	16.1	16.1	12.0	11.5	4.6	8.3	.5	2.8	.5	1.5	0	.7
41– 50.....	13.0	9.5	13.5	13.5	8.9	10.6	2.4	5.2	.8	.9	0	.3
51– 60.....	11.2	5.6	12.4	17.6	10.1	14.1	5.0	9.2	2.2	3.9	.4	1.0
61– 70.....	9.6	3.3	14.6	13.2	17.8	16.2	8.4	11.2	3.8	12.2	1.5	5.1
71– 80.....	4.7	2.0	17.1	9.8	21.0	18.3	19.6	20.4	13.4	16.6	6.9	5.5
81– 90.....	1.8	.3	12.7	8.4	19.6	20.0	30.6	20.4	32.0	33.8	21.7	33.1
91–100.....	.7	0	5.0	2.4	13.2	6.2	33.1	24.4	47.1	30.0	69.7	54.3
Medians...	35.80	25.27	60.70	51.75	73.10	67.67	84.90	79.36	90.50	84.86	94.68	91.96

difference will occasion no surprise. We should do well also to bear in mind not only that schools do vary greatly, but that in this particular instance there was a constant factor tending to lower the ratings. In the supplementary test the dictation was given by a stranger; in the original test by the class teacher, guided by printed directions. Leaving out of account any sugges-

tion of unfair methods on the part of teachers for the purpose of making a showing, this fact is quite sufficient to account for a falling off of results in schools VI and VII without supposing them to be much, if any, inferior to the others in the ability of their children to spell. The teacher has but one class to examine and she takes her time. She doubtless takes full advantage of the direction permitting the reading of a sentence "in whole or in part as many times as may be necessary to secure its full comprehension." She knows her class. The peculiarities or defects of pupils are her daily concern and she modifies her appeal accordingly. The children are at ease in her presence. They know her voice and manner of speaking; and they more readily understand her than they would another. In these matters they are placed at a disadvantage when examined by a stranger.

We should expect this disadvantage to be most evident in the lower classes; and an inspection of the medians of Table XXVI shows how strikingly this is true. The children of the 3d grade fell off 10.5%; of the 4th, 9%; of the 5th, 6th, and 7th approximately 5.5%; and of the 8th only 2.5%. The force, whatever its nature, tending to depress the results at schools VI and VII was clearly operative to a greater degree in the lower classes and to a much less degree in higher classes. The effect of a change from the class teacher to a stranger as examiner would be expected to bring about results of just such a nature.

But the obvious advantage of having the same examiner for every class (in this case there were two working together) is that however the results in general may be lowered, there is a better opportunity to compare class with class or grade with grade, or school with school.

The real reason why this supplementary testing was undertaken was to verify the median intervals that had been derived from the original testing. The fact of classes being examined by the same persons is of great value for this purpose. If, with this factor of the examiner made constant, we find that these median distances in spite of a falling off of grade performances remain substantially the same, we shall be justified in feeling that our former results are reasonably reliable.

TABLE XXVII

NUMBER AND PER CENT OF PUPILS IN EACH GRADE WHOSE ABILITY EQUALLED OR EXCEEDED THAT OF THE MEDIAN PUPIL IN EVERY OTHER GRADE, WITH THE P.E. VALUES CORRESPONDING TO EACH PER CENT. SCHOOLS VI AND VII COMBINED WITH SCHOOLS II, III, IV AND V

		3d Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade
3d grade...	No.		110	38	12	4	0
N=749....	%		14.7	5.1	1.6	.5	0
	P.E.		1.55	2.425	3.18	3.82	?
4th grade..	No.	619		223	90	45	16
N=763....	%	81.1		29.2	11.8	5.9	2.1
	P.E.	—1.31		.81	1.76	2.32	3.01
5th grade..	No.	758	584		227	121	47
N=805....	%	94.2	72.5		28.2	15.0	5.8
	P.E.	—2.33	—.89		.86	1.54	2.33
6th grade..	No.	655	597	511		241	112
N=668....	%	98.1	89.4	76.5		36.1	16.8
	P.E.	—3.08	—1.85	—1.07		.55	1.43
7th grade..	No.	698	678	615	482		188
N=702....	%	99.43	96.6	87.6	68.7		26.8
	P.E.	—3.75	—2.71	—1.71	—.72		.92
8th grade..	No.	570	565	544	502	428	
N=570....	%	100.0	99.12	95.44	88.1	75.1	
	P.E.	?	—3.52	—2.50	—1.75	—1.00	

In Table XXVII we give the number and per cent of pupils who equal or surpass the medians of other grades than their own with the corresponding P.E. values. In this table are combined the children who were examined at schools VI and VII with those examined at schools II, III, IV and V. It is to be compared with Table XV (page 36). Table XXVIII gives the median distances as derived from the P.E. values of Table XXVII. Compare with Table XVI (page 39). The average distances, 1.37, .87, .90, .66, and .86, are to be compared with the values for the same distances as derived on page 39, namely (correct to 2 decimal places) 1.35, .84, 1.05, .66, and .91. Only in the case of $M_{5\text{-}6}$ is there an important difference. The average difference, including that of $M_{5\text{-}6}$, is .05; excluding it, the average difference is but .025. The entire range from M_3 to M_8 is found

to be 4.66 as compared with 4.81 according to the primary testing. These correspondences are, we feel, quite close enough to establish the essential reliability of our original figures.

TABLE XXVIII

MEDIAN DISTANCES DERIVED FROM THE P.E. VALUES OF TABLE XXVII

	M_{3-4}	M_{4-5}	M_{5-6}	M_{6-7}	M_{7-8}
	1.55 (direct)	.87 (M_{3-5}—M_{3-4})	.76 (M_{3-6}—M_{3-5})	.64 (M_{3-7}—M_{3-6})	? (M_{3-8}—M_{3-7})
	1.61 (M_{3-5}—M_{4-5})	.81 (direct)	.95 (M_{4-6}—M_{4-5})	.56 (M_{4-7}—M_{4-6})	.69 (M_{4-8}—M_{4-7})
	1.42 (M_{3-6}—M_{4-6})	.90 (M_{4-6}—M_{5-6})	.86 (direct)	.68 (M_{5-7}—M_{5-6})	.79 (M_{5-8}—M_{5-7})
	1.50 (M_{3-7}—M_{4-7})	.78 (M_{4-7}—M_{5-7})	.99 (M_{5-7}—M_{6-7})	.55 (direct)	.88 (M_{6-8}—M_{6-7})
	? (M_{3-8}—M_{4-8})	.68 (M_{4-8}—M_{5-8})	.90 (M_{5-8}—M_{6-8})	.51 (M_{6-8}—M_{7-8})	.92 (direct)
	1.31 (direct)	1.02 (M_{5-3}—M_{4-3})	.75 (M_{6-3}—M_{5-3})	.67 (M_{7-3}—M_{6-3})	? (M_{8-3}—M_{7-3})
	1.44 (M_{5-3}—M_{5-4})	.89 (direct)	.96 (M_{6-4}—M_{5-4})	.86 (M_{7-4}—M_{6-4})	.81 (M_{8-4}—M_{7-4})
	1.23 (M_{6-3}—M_{6-4})	.78 (M_{6-4}—M_{6-5})	1.07 (direct)	.64 (M_{7-5}—M_{6-5})	.79 (M_{8-5}—M_{7-5})
	1.04 (M_{7-3}—M_{7-4})	1.00 (M_{7-4}—M_{7-5})	.99 (M_{7-5}—M_{7-6})	.72 (direct)	1.03 (M_{8-6}—M_{7-6})
	? (M_{8-3}—M_{8-4})	1.02 (M_{8-4}—M_{8-5})	.75 (M_{8-5}—M_{8-6})	.75 (M_{8-6}—M_{8-7})	1.00 (direct)
Average	1.37	.87	.90	.66	.86

§ 17. *Arrangement of the Words of the Preferred List by Teachers' Judgments*

A certain order of difficulty of the 50 words of the Preferred List having been determined as the result of testing in five schools and to a certain extent verified by a record from two other schools (Table XIX, and Fig. 20), a comparison of the result with an arrangement of the same words based on the judgment of teachers becomes interesting. It has an important

bearing on the whole spelling situation because it is by individual judgment as to the difficulty of words that lists are made up and graded for classes, schools, or systems of schools. How far such grading is reliable may be gathered by finding out to what extent individual judgment squares with the results of actual testing.

For this purpose the 50 words of the Preferred List were arranged alphabetically and distributed to a number of teachers. They were asked to arrange the words in what they judged to be their order of difficulty for children to spell, beginning with the easiest. They were particularly requested to do the work without consulting any one. Two hundred arrangements were secured. They differed widely—so widely that whatever may be the value of a consensus of many individuals, the trustworthiness of the judgment of a single teacher appears to be almost of no value. It may be good and it may be bad; and it is about as likely to be the one as the other. With one notable exception, the agreement with the record was closer for the easiest and hardest words than for those of medium difficulty. This might have been expected from the fact that, as shown in Fig. 20, the words at the ends of the scale differ by larger amounts than do those nearer the middle. The exception referred to is the word "too," which, although it was in every school the hardest word to spell, was by more than a fourth of the teachers judged to be actually the easiest, or next to the easiest, in the list. The deviation of individual judgments from the record is shown by figures taken from a random sampling of the two hundred teachers' arrangements. Five such arrangements being taken by chance from the whole number, proved to be those of teachers No. 7, No. 88, No. 109, No. 134, and No. 178. They ranked the 5th, 10th, 15th, 20th, ... 50th words (record) as follows:

Record	5, 10, 15, 20, 25, 30, 35, 40, 45, 50	
No. 7	22, 15, 1, 31, 34, 33, 38, 47, 14, 10	$r=+0.09$
No. 88	15, 9, 4, 41, 13, 24, 33, 35, 20, 1	$r=+0.15$
No. 109	9, 11, 18, 39, 35, 36, 22, 25, 29, 10	$r=+0.15$
No. 134	24, 22, 23, 20, 6, 44, 7, 42, 35, 11	$r=-0.03$
No. 178	21, 2, 7, 22, 14, 29, 49, 38, 28, 1	$r=+0.2$

It is, however, to be expected that when a large number of judgments are taken together, wide disagreements with a true arrangement will tend to disappear, and a resultant will be obtained that may be expected to be closer to the facts than any single judgment would ever be likely to be. The statistical treatment of the 200 judgments was based on the theorem, "Differences that are equally often noticed are equal, unless the differences are either always or never noticed." It is an abbreviation of the method used by Professor Thorndike in deriving his scale for Handwriting and by Dr. Hillegas in his similar work for English Composition. We have not felt the necessity of making comparisons of the judgment of each word with that of every other word, because the nature of our material has enabled us to derive our scale by a more direct method. Since we are here concerned with a comparison only, we have been content to proceed as follows: The record shows "only" to be easier than "even." What per cent of the individuals who arranged the words for difficulty so judged? As between "even" and "smoke"; "smoke" and "chicken"; "chicken" and "front"; etc., what per cent of the judgments indicate that the first word of each pair is easier than the second, as the record shows? The following is found to be true for the first six words:

"only"	was	judged	easier	than	"even"	by	38.5%	of	the	judges.
"even"	"	"	"	"	"smoke"	"	67.5%	"	"	"
"smoke"	"	"	"	"	"chicken"	"	72.0%	"	"	"
"chicken"	"	"	"	"	"front"	"	48.0%	"	"	"
"front"	"	"	"	"	"another"	"	53.0%	"	"	"

Now when the difficulty of a word is judged by a very great number of judges some will overestimate its difficulty, others will underestimate it. Those who make small errors will be more numerous than those who make large errors. The frequency of the judgments will take the form of the curve of the probability integral whose base-line represents the amounts of difficulty which the word in question is judged to have. The point on the base-line which corresponds to the greatest frequency of judgments represents the central tendency of the judges in rating the word. It is therefore the point which represents the difficulty of the word as determined by individual judgments. Two or more words may be compared for difficulty if we know the per

cent of judges who rate one word easier (or harder) than the other.

Fig. 22 shows the curves for the first six words arranged to show the per cent of "easier" judgments noted above. The curve for "even" is so placed that 38.5% of its area is to the right of *YO*—the median axis of the curve for "only" when that

TABLE XXIX

COMPARISON OF RESULTS BY TEACHERS' JUDGMENTS AND BY THE RECORD. PREFERRED LIST

Word-Number (scale)	Word	% of times each word was judged easier than the following word	Rank		Word-Number (scale)	Word	% of times each word was judged easier than the following word	Rank	
			By Teachers' Judgments	By the Record				By Teachers' Judgments	By the Record
3	only.....	38.5	2	1	14	minute....	25.5	27	26
1	even.....	67.5	1	2	7	pear.....	51.0	15	27
4	smoke.....	72.0	3	3	40	towel.....	69.0	16	28
28	chicken....	48.0	7	4	38	tobacco...	56.0	25.5	29
5	front.....	53.0	6	5	43	whole.....	15.5	28	30
9	another...	39.5	8	6	13	button....	79.5	9	31
2	lesson.....	86.0	4	7	17	janitor....	81.0	23	32
8	bought....	29.5	18	8	33	quarrel....	39.5	44	33
11	pretty.....	27.5	10	9	44	against....	43.5	40	34
16	nails.....	89.0	11	10	30	circus.....	77.0	35	35
46	butcher...	65.5	19	11	20	sword.....	44.5	48	36.
41	Tuesday...	19.5	29	12	23	whistle....	21.5	46	36.5
6	sure.....	74.5	13	13	19	stopping...	87.0	30	38
45	answer....	10.0	21.5	14	24	carriage...	29.0	49	39
25	nor.....	90.0	5	15.5	47	guess.....	54.0	43	40
49	raise.....	69.0	21.5	15.5	37	telephone..	27.0	45	41
15	cousin....	53.0	37	17.5	29	choose....	56.0	31	42
50	beautiful..	41.5	39	17.5	36	telegram...	54.5	36	43
22	touch.....	29.5	33	19	18	saucer.....	58.0	38	44
21	freeze.....	31.5	20	20	34	saucy.....	35.5	41	45
10	forty.....	75.0	12	21.5	26	already....	91.5	34	46
48	instead....	33.0	24	21.5	32	pigeons....	16.0	50	47
12	wear.....	44.0	17	23	27	beginning..	64.0	42	48
35	tailor.....	72.5	14	24.5	31	grease.....	21.0	47	49
42	tying.....	55.0	25.5	24.5	39	too.....	....	32	50

axis is produced; 67.5% of the curve for "smoke" is to the right of Y^1O^1 produced; 72% of the curve for "chicken" is to the right of Y^2O^2 produced; and so on. $H\,O^1$ is the difference in difficulty between "only" and "even"; $K\,O^2$, between "even" and "smoke"; $P\,O^3$, between "smoke" and "chicken"; $S\,O^4$, between "chicken" and "front"; and $T\,O^5$, between "front" and "another." When less than 50% of the judges regard the first of a pair of words as easier, the second is, of

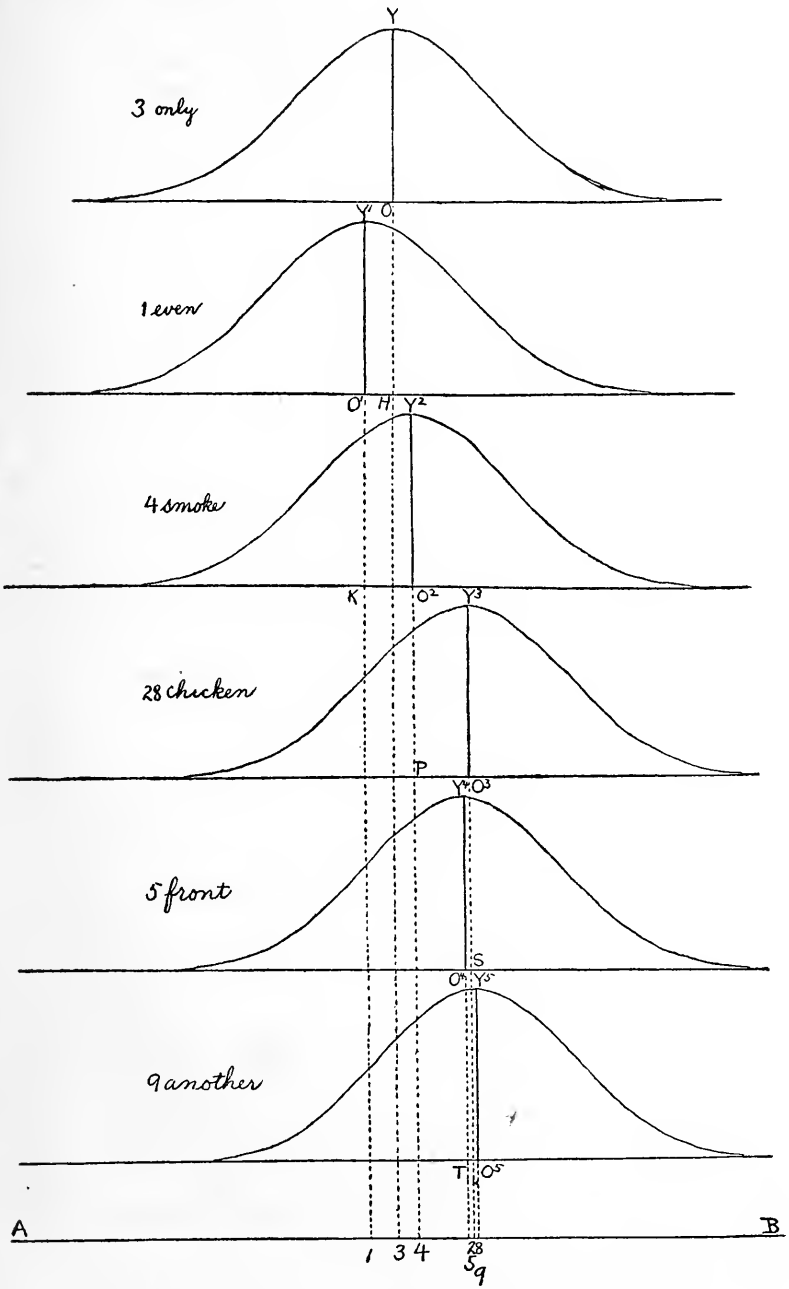


FIG. 22. Diagram showing difference in difficulty between words by Teachers' Judgments.

course, judged the easier, and the difference in difficulty is a negative one. Such is the case in the *only-even* and *chicken-front* pairs. The scaling of these six words is diagrammatically shown in Fig. 22 by producing each median vertical to meet the line *AB*. It will be seen that the order of difficulty is not the same as the order obtained by testing.

Table XXIX shows for the entire 50 words the per cent of teachers who judged each word easier than the following word and the rank of each word for difficulty by record and by teachers' judgments. In spite of the fact that the opinion of single teachers is so unreliable, the combined judgments of a group as large as 200 yield an arrangement which agrees closely enough with the arrangement by record to confirm and support the latter in no small degree. The correlation, by the 'foot-rule' method, is found to be 0.79, which may quite properly be regarded as satisfactory.

It may be worth while to point out, however, that in practice the selection and arrangement of words for teaching are not the work of a large number of individuals. These things are usually done by a single teacher for a class or by a text-book writer for as many classes as use his book. Not two hundred, probably not even ten, persons judge as to the selection and arrangement of the words in the lists now used in most schools. The length of such lists, moreover, would seem to preclude the possibility of a satisfactory judgment as to difficulty by individuals. Probably if our own list of 50 words had been shorter, the teachers would have worked more accurately. The several thousand words in a spelling-book certainly constitute a list about which there may be expected to be wide and numerous disagreements. We alluded in Section 3 above to our attempt to secure for school use a 5000-word vocabulary graded by years and based upon the agreements of five spelling-books. This task proved to be very difficult precisely because of the total absence of agreement as to grading in the case of hundreds of words. One speller would assign words to the third grade which another would put in the sixth, seventh, or eighth. Gradings three, four, and even five years apart occurred with remarkable frequency.

The obvious way (and the necessary way, it would seem) to grade words for difficulty is not by some one's opinion of how

hard they are, but by actually "trying them out." In matters of handwriting and composition the judgment of individuals is all-important, because merit in either is precisely a matter of judgment. One sample is better than another only because competent persons think so. On the contrary, one word is harder to spell than another not because we think so, but because more people misspell the one than the other, or because it takes more time to learn to spell the one than the other. It is strange, therefore, that no spelling-book has yet appeared based upon a study of how frequently children misspell the words of which it is composed. In fact no study of spelling, that we know of, has done more than obtain individual ratings of pupils, based on the tacit assumption that the words used are all equally difficult to spell. No investigation has been thought necessary of the words themselves. The results of this section, although by no means thoroughly worked out, sufficiently indicate the present unreliability of individual judgment with regard to words, unless the list is very short and the judgments very numerous. It is quite possible that at some later time, after studies of words based on actual tests have been frequently made, our judgment of word difficulties may be greatly improved. Our opinion as to how hard words are might then become a valuable supplement to the conclusions of the investigator.

§ 18. *Rice Sentence Test. Easy 50-Word Test*

During the middle week of the fall term of 1912, the sentence test used by Rice—and afterwards by Cornman—was dictated to 1984 pupils in schools II, III, and VIII. Children of the 4th and 5th grades wrote sentences containing 50 words; those of the 6th, 7th, and 8th grades wrote 41 of the same words together with 36 additional words—77 in all. The entire test follows:

Rice Sentence Test

While *running* he *slipped.* I *listened* to his *queer speech,* but I did not *believe* any of it. The *weather* is *changeable.* His loud *whistling frightened* me. He is *always changing* his mind. His

chain was *loose.* She was *baking* cake. I have a *piece* of it. Did you *receive* my letter? I heard the *laughter* in the *distance.* Why did you *choose* that *strange picture?* *Because I *thought* I liked it. It is my *purpose* to *learn.* Did you *lose* your *almanac?* I gave it to my *neighbor.* *I was *writing* in my *language* book. Some children are not *careful enough.* Was it *necessary* to keep me *waiting* so long? Do not *disappoint* me so *often.* I have *covered* the *mixture.* He is *getting better.* *A *feather* is *light.* Do not *deceive* me. I am *driving* a new horse. *Is the *surface* of your desk *rough* or *smooth?* The children were *hopping.* This is *certainly* true. I was very *grateful* for my *elegant present.* If we have *patience* we will *succeed.* He met with a *severe accident. Sometimes* children are not *sensible.* You had no *business* to *answer* him. You are not *sweeping properly.* Your reading shows *improvement.* The ride was very *fatiguing.* I am very *anxious* to hear the news. I *appreciate* your kindness I *assure* you. I cannot *imagine* a more *peculiar character.* I *guarantee* the book will meet with your *approval. Intelligent* persons learn by *experience.* The peach is *delicious.* I *realize* the *importance* of the *occasion.* Every rule has *exceptions.* He is *thoroughly conscientious; therefore* I do trust him. The elevator is *ascending.* Too much *praise* is not *wholesome.*

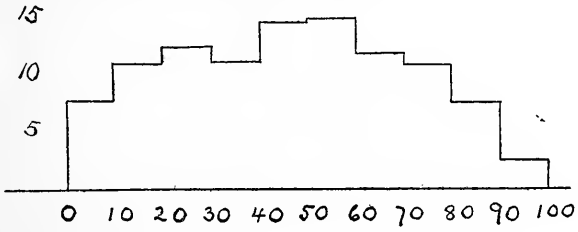
TABLE XXX

DISTRIBUTION OF INDIVIDUAL RATINGS. RICE SENTENCE TEST.

Per Cent Correct	4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
	No.	%	No.	%	No.	%	No.	%	No.	%
0– 10	38	7.5	3	.6	2	.5	0	0	0	0
11– 20	54	10.7	13	2.8	13	3.3	3	.8	0	0
21– 30	61	12.0	18	3.8	21	5.3	5	1.4	0	0
31– 40	55	10.8	49	10.4	46	11.6	20	5.4	0	0
41– 50	72	14.2	62	13.2	53	13.4	32	8.7	2	.8
51– 60	74	14.6	76	16.2	80	20.2	52	14.2	15	6.1
61– 70	59	11.6	76	16.2	70	17.7	53	14.4	37	15.2
71– 80	54	10.7	91	19.4	62	15.7	77	21.0	61	25.0
81– 90	31	6.1	65	13.8	40	10.1	95	25.9	96	39.3
91–100	9	1.8	17	3.6	9	2.3	30	8.2	33	13.5
Totals	507		470		396		367		244	
Medians		48.17		64.13		58.57		73.86		82.07

The 4th- and 5th-year test ends with " This is *certainly* true." The test for the upper grades comprises all the sentences except the four marked with an asterisk. The test words are italicized.

The principal object in giving this test was to obtain scores for a new series of words by which the grade and general scales (Figs. 14-20) could be supplemented and extended. The papers, however, were also rated for individual performances. Subject



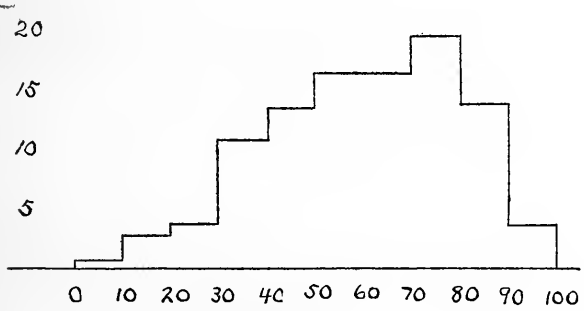


Fig. 23. Relative frequencies of different percentages correct, 4th grade; Rice Sentence Test; Table XXX.

Fig. 24. Same as Fig. 23, but for 5th grade.

to the limitations of regarding all words as equal, the results on this basis may be used to supplement those of Rice and Cornman. Rice gives nothing but grade averages, and Cornman gives the same for two tests after an interval of one year. We shall continue to use the median as a measure of central tendency and shall give a distribution of pupils' ratings for each grade. Table XXX gives the distribution by groups of ten with totals and medians. Fig. 23 and Fig. 24 show graphically the distribution

for the 4th and 5th grades. Fig. 8, Fig. 9, and Fig. 10 (page 33) give the same for 6th, 7th, and 8th grades.

These medians indicate a performance for these schools poorer than Rice indicates for most of his schools and much poorer than Cornman's results for the two schools which he tested. We cannot account for this because neither of these investigators tells how he rated the pupils' papers. In our own testing omitted and illegible words were counted as wrong. It is probable that the manner in which Rice chose his schools would give him those in which better than average work was being done; while Cornman's two schools were without doubt devoting an unusual amount of attention to spelling under his personal guidance. Such being the case, it is quite possible that the results here given more nearly approach typical conditions, than do those of either of these writers.

TABLE XXXI

PER CENT CORRECT FOR EACH WORD IN EACH GRADE WITH CORRESPONDING P.E. VALUES. RICE SENTENCE TEST. SEE FIG. 25

No. of Word	Word	4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
		%	P.E.	%	P.E.	%	P.E.	%	P.E.	%	P.E.
1	running	48.0	+ .074	66.0	— .612	76.8	—1.086	85.0	—1.537	93.4	—2.234
2	slipped	30.0	+ .777	34.8	+ .579	42.9	+ .265	51.8	— .067	70.9	— .816
3	listened	29.6	+ .794	40.4	+ .360	53.5	— .130	69.8	— .769	86.9	—1.663
4	queer	56.9	— .258	58.8	— .330	77.3	—1.110	79.0	—1.196	87.3	—1.692
5	speech	45.3	+ .175	41.4	+ .322	72.0	— .864	77.1	—1.101	80.7	—1.286
6	believe	37.2	+ .484	49.7	+ .011	64.4	— .547	62.1	— .457	76.6	—1.076
7	weather	70.9	— .816	57.5	— .280	82.8	—1.403	88.0	—1.742	92.2	—2.103
8	changeable	27.7	+ .878	31.3	+ .723	46.7	+ .123	66.8	— .604	65.6	— .596
9	whistling	27.3	+ .895	40.0	+ .376	49.0	+ .037	68.7	— .723	74.2	— .963
10	frightened	17.8	+1.368	42.7	+ .273	55.6	— .209	71.4	— .838	85.7	—1.582
11	always	53.8	— .141	68.8	— .727	78.5	—1.170	88.6	—1.788	95.5	—2.514
12	changing	58.5	— .318	69.2	— .744	74.5	— .974	89.6	—1.867	91.4	—2.035
13	chain	51.8	— .067	59.8	— .368	75.3	—1.014	88.0	—1.742	95.9	—2.579
14	loose	24.7	+1.014	49.1	+ .033	45.2	+ .179	63.2	— .500	81.6	—1.335
15	baking	63.4	— .508	75.7	—1.033	83.6	—1.450	93.5	—2.245	97.5	—2.905
16	piece	58.5	— .318	62.2	— .461	69.9	— .773	83.7	—1.456	90.6	—1.953
17	receive	21.1	+1.191	51.7	— .063	59.8	— .368	62.1	— .457	80.7	—1.286
18	laughter	59.9	— .372	71.4	— .838	75.5	—1.024	88.8	—1.803	96.3	—2.648
19	distance	35.6	+ .547	67.2	— .660	75.8	—1.038	88.0	—1.742	97.5	—2.905
20	choose	41.7	+ .311	46.3	+ .138	56.8	— .254	83.1	—1.421	85.7	—1.582
21	strange	57.7	— .288	74.2	— .963	86.9	—1.663	93.5	—2.245	92.6	—2.145
22	picture	69.6	— .761	87.5	—1.706	94.4	—2.357	97.5	—2.905	98.8	—3.346
23	because	66.2	— .620	83.9	—1.459						
24	thought	58.7	— .326	72.4	— .882						
25	purpose	21.7	+1.160	47.3	+ .100	66.9	— .648	74.7	— .986	92.6	—2.145
26	learn	70.6	— .803	84.9	—1.531	93.2	—2.211	95.9	—2.579	99.6	—3.938
27	lose	46.4	+ .134	53.1	— .119	56.8	— .254	60.0	— .376	55.7	— .213
28	almanac	10.1	+1.892	21.5	+1.170	38.6	+ .430	58.6	— .322	72.1	— .869
29	neighbor	27.5	+ .886	66.6	— .636	65.2	— .579	85.0	—1.537	93.4	—2.234
30	writing	56.3	— .235	74.0	— .954						

TABLE XXXI

(*Continued*)

No. of Word	Word	4th Grade		5th Grade		6th Grade		7th Grade		8th Grade	
		%	P.E.	%	P.E.	%	P.E.	%	P.E.	%	P.E.
31	language	40.3	+ .364	62.8	— .484						
32	careful	54.3	— .160	58.6	— .322	68.9	— .731	85.8	—1.589	88.1	—1.749
33	enough	54.9	— .183	68.0	— .693	80.3	—1.264	91.0	—1.988	98.4	—3.182
34	necessary	4.5	+2.514	21.5	+1.170	42.7	+ .273	37.6	+ .468	61.5	— .434
35	waiting	55.9	— .220	66.8	— .604	82.3	—1.374	89.9	—1.892	92.2	—2.103
36	disappoint	11.7	+1.757	27.4	+ .891	34.6	+ .588	32.4	+ .657	38.9	+ .418
37	often	51.6	— .059	57.5	— .280	75.8	—1.038	87.2	—1.685	92.2	—2.103
38	covered	42.1	+ .296	62.6	— .476	77.5	—1.120	90.2	—1.918	97.1	—2.811
39	mixture	33.6	+ .628	62.6	— .476	83.3	—1.432	91.0	—1.988	97.1	—2.811
40	getting	57.5	— .280	74.4	— .972	87.4	—1.699	94.6	—2.384	97.5	—2.905
41	better	80.6	—1.279	91.8	—2.064	94.9	—2.425	98.6	—3.258	100.0	?
42	feather	77.1	—1.101	84.1	—1.481						
43	light	77.5	—1.120	90.5	—1.944						
44	deceive	18.4	+1.335	46.3	+ .138	53.5	— .130	54.8	— .179	79.5	—1.222
45	driving	59.7	— .364	77.1	—1.101	88.1	—1.749	65.7	—2.177	98.8	—3.346
46	surface	48.4	+ .059	79.1	—1.201						
47	rough	64.2	— .539	69.8	— .769						
48	smooth	47.2	+ .104	51.3	— .048						
49	hopping	58.1	— .303	58.1	— .303	68.4	— .710	81.2	—1.313	89.3	—1.843
50	certainly	16.8	+1.427	36.0	+ .531	57.1	— .265	79.0	—1.196	91.0	—1.988
51	grateful					39.1	+ .410	58.6	— .322	61.9	— .449
52	elegant					53.5	— .130	65.7	— .600	69.3	— .748
53	present					69.7	— .765	79.0	—1.196	91.4	—2.035
54	patience					43.4	+ .246	63.0	— .492	80.7	—1.286
55	succeed					53.0	— .112	70.8	— .812	80.7	—1.286
56	severe					40.9	+ .341	61.3	— .426	70.9	—. 816
57	accident					45.5	+ .168	68.9	— .731	85.2	—1.549
58	sometimes					52.5	— .093	67.3	— .665	82.8	—1.403
59	sensible					34.3	+ .600	55.0	— .187	65.2	— .579
60	business					46.0	+ .149	53.7	— .138	68.4	— .710
61	answer					74.0	— .954	86.9	—1.663	93.4	—2.234
62	sweeping					87.4	—1.699	92.1	—2.093	94.7	—2.397
63	properly					61.1	— .418	73.0	— .909	86.5	—1.636
64	improvement					59.6	— .360	69.5	— .756	86.5	—1.636
65	fatiguing					12.6	+1.699	25.3	+ .986	31.1	+ .731
66	anxious					49.0	+ .037	66.2	— .620	84.0	—1.475
67	appreciate					31.8	+ .702	49.0	+ .037	74.6	— .982
68	assure					58.1	— .303	68.9	— .731	86.1	—1.609
69	imagine					33.6	+ .628	47.7	+ .085	66.4	— .628
70	peculiar					24.0	+1.047	46.3	+ .138	56.1	— .228
71	character					40.2	+ .368	47.1	+ .108	78.7	—1.181
72	guarantee					11.6	+1.772	19.9	+1.253	25.8	+ .963
73	approval					38.1	+ .449	56.9	—1.258	75.4	—1.019
74	intelligent					37.1	+ .448	43.6	+ .239	50.4	— .015
75	experience					44.4	+ .209	63.5	— .512	68.9	— .731
76	delicious					31.3	+ .723	61.6	— .437	85.2	—1.549
77	realize					53.5	— .130	65.7	— .600	73.4	— .927
78	importance					47.5	+ .093	73.3	— .922	81.6	—1.335
79	occasion					34.8	+ .579	44.4	+ .209	49.6	+ .015
80	exceptions					48.2	+ .067	57.2	— .269	76.2	—1.057
81	thoroughly					18.7	+1.318	31.1	+ .731	53.3	— .123
82	conscientious					.3	+4.167	1.6	+3.171	19.7	+1.264
83	therefore					36.4	+ .516	62.4	— .468	79.9	—1.243
84	ascending					37.6	+ .468	52.0	— .074	55.7	— .213
85	praise					69.0	— .735	78.2	—1.155	95.9	—2.579
86	wholesome					56.3	— .235	74.7	— .986	86.1	—1.609

With respect to the ratings of words, Table XXXI gives for each grade the per cent of correct spellings and the P.E. values calculated from the grade medians, assuming a normal distribution. Fig. 25 (insert) shows the lines the same words arranged on a linear scale for grades 4, 5, 6, 7, and 8. Above the lines the arrangement of the words of the Preferred List is given. This latter is a repetition of the scales of Figures 15, 16, 17, 18 and 19 (p. 44). It will at once be seen that the former scales, obtained by using the Preferred List, have been filled in and have been extended much further to the right.

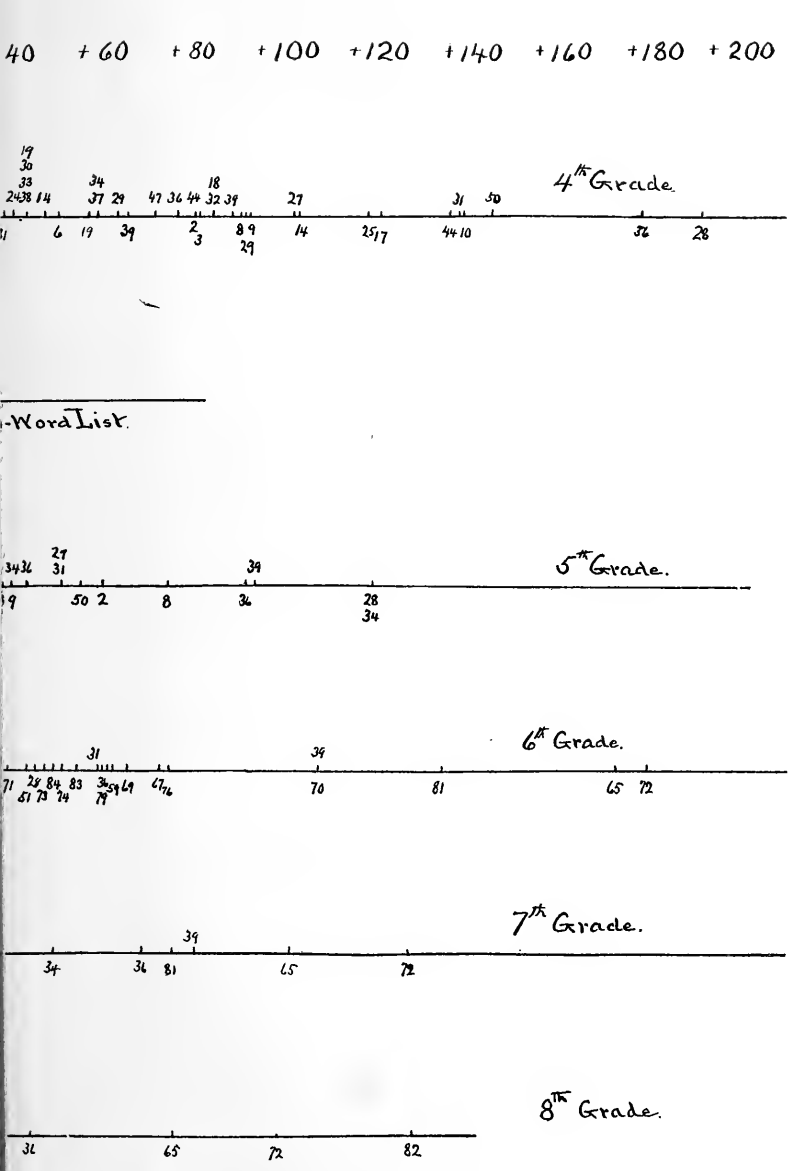
Just as the more difficult words of the Rice Test may be used to extend the scales to the right, so the easier words of the Easy 50-Word Test may be used to extend it towards the left in certain grades. In the grades of the second school year the latter were the only words used. Although the primary object in giving the Easy 50-Word Test was to enable us to give a position to the zero-point, and although for this purpose the ratings of individual pupils were sufficient, nevertheless the per cent of correct spellings for each word in each grade (*2a, 2b,* 3d, and 4th) was also calculated.

Table XXXII gives these per cents and the corresponding P.E. values. It will be noted that there are six words (*even, only, pretty, sure, touch,* and *front*) that are common to this list and to the Preferred List.

TABLE XXXII

PER CENT CORRECT FOR EACH WORD IN EACH GRADE WITH CORRESPONDING P.E. VALUES. EASY 50-WORD TEST

No. of Word	Word	2a Grade 175 Pupils		2b Grade 169 Pupils		3d Grade 168 Pupils		4th Grade 316 Pupils	
		%	P.E.	%	P.E.	%	P.E.	%	P.E.
1	you.....	52.6	— .097	71.6	— .847	83.9	—1.459	93.4	—2.234
2	will.....	46.9	+ .119	83.4	—1.438	96.4	—2.667	99.1	—3.506
3	hear.....	37.1	+ .489	53.3	— .123	60.1	— .380	79.4	—1.217
4	him.....	25.7	+ .968	71.6	— .847	81.0	—1.302	97.5	—2.905
5	coming..	9.7	+1.926	33.7	+ .624	57.1	— .265	83.5	—1.444
6	he.....	62.3	— .464	88.7	—1.795	98.2	—3.111	99.7	—4.083
7	is.....	65.7	— .600	94.7	—2.397	96.4	—2.667	98.4	—3.182
8	on.....	57.7	— .288	88.7	—1.795	94.6	—2.384	97.8	—2.986
9	the.....	65.7	— .600	94.7	—2.397	94.6	—2.384	97.8	—2.986
10	road....	4.6	+2.498	20.7	+1.211	53.6	— .134	82.6	—1.391



he Easy 50-Word List, Table XXXII. These scales are not absolutely
2 of the Rice List at +317. For the lists referred to see Appendix II.

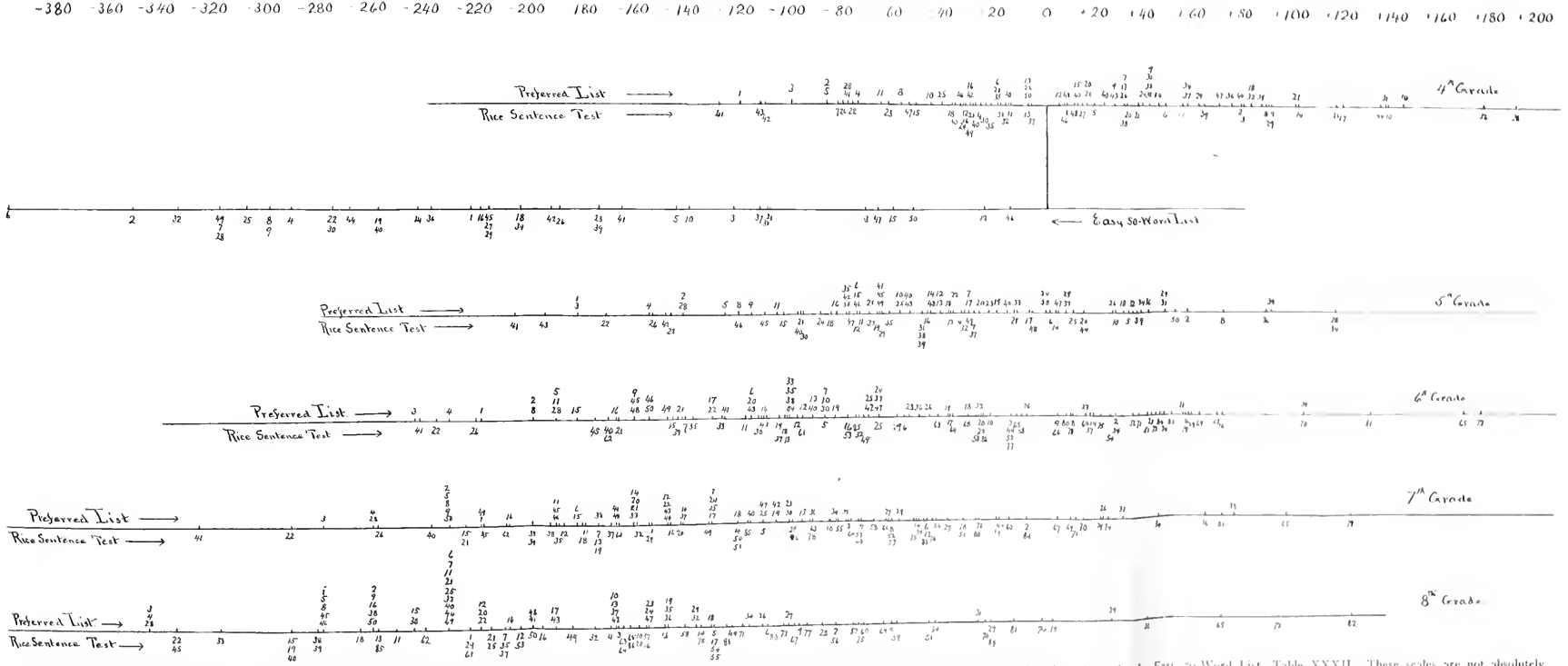


FIG. 25 Completed Grade Scale. Grades 4 to 8. Combining the Rice Sentence List, Table XXXI, with the Preferred List, Table XVII, to which is also added, for the 4th grade, the Easy 50 Word List, Table XXXII. These scales are not absolutely complete. The 4th grade scale should have word No. 34 of the Rice List at +255; the 6th grade scale should have word 82 of the Rice List at +415; the 7th grade scale should have 82 of the Rice List at +317. For the lists referred to see Appendix II

TABLE XXXII

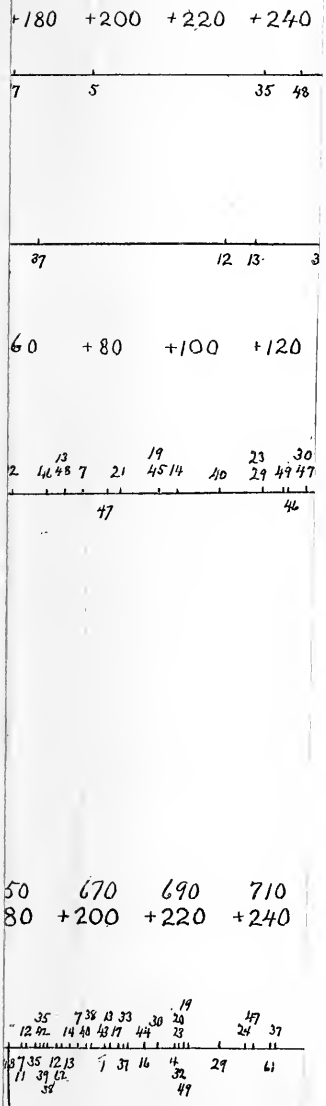
(*Continued*)

No. of Word	Word	2a Grade 175 Pupils		2b Grade 169 Pupils		3d Grade 168 Pupils		4th Grade 316 Pupils	
		%	P.E.	%	P.E.	%	P.E.	%	P.E.
11	and.....	51.4	— .052	85.2	—1.549	97.0	—2.789	98.4	—3.182
12	almost...	.6	+3.725	6.5	+2.245	42.9	+ .265	56.3	— .235
13	sure.....	1.1	+3.392	5.9	+2.321	48.8	+ .044	61.1	— .418
14	to.....	48.0	+ .074	78.7	—1.181	93.5	—2.245	94.9	—2.425
15	pass.....	1.1	+3.392	27.8	+ .873	51.2	— .044	65.8	— .603
16	in.....	45.7	+ .160	91.1	—1.997	92.9	—2.177	93.0	—2.188
17	front....	12.0	+1.742	27.8	+ .873	48.2	+ .067	67.1	— .656
18	of.....	25.7	+ .968	64.5	— .551	82.1	—1.363	91.5	—2.035
19	me.....	54.3	— .160	84.6	—1.512	88.7	—1.795	95.9	—2.579
20	I.....	45.1	+ .183	93.5	—2.245	95.2	—2.468	100.0	?
21	send....	22.3	+1.130	25.4	+ .982	48.2	+ .067	76.6	—1.076
22	for.....	34.3	+ .600	78.7	—1.181	93.5	—2.245	96.8	—2.746
23	every....	3.4	+2.706	39.6	+ .391	66.7	— .640	88.0	—1.742
24	day.....	36.6	+ .508	89.3	—1.843	94.0	—2.305	100.0	?
25	go.....	45.1	+ .183	87.0	—1.670	89.3	—1.843	98.1	—3.077
26	into.....	34.3	+ .600	71.0	— .820	76.2	—1.057	89.9	—1.892
27	school...	16.0	+1.475	60.4	— .371	81.5	—1.329	92.7	—2.155
28	but.....	13.1	+2.767	58.0	— .299	83.3	—1.432	98.4	—3.182
29	do.....	40.1	+ .372	74.0	— .954	85.1	—1.543	97.2	—2.155
30	not.....	33.7	+ .624	69.2	— .744	86.9	—1.663	96.8	—2.746
31	touch...	.6	+3.725	4.7	+2.483	29.8	+ .786	54.4	— .164
32	table....	1.7	+3.146	13.0	+1.670	53.6	— .239	94.3	—2.344
33	also.....	2.9	+2.811	15.4	+1.512	35.7	+ .543	68.4	— .710
34	has.....	36.0	+ .531	65.1	— .575	69.6	— .761	91.5	—2.035
35	only.....	5.7	+2.344	16.6	+1.438	53.0	— .112	70.6	— .803
36	one.....	34.9	+ .575	71.0	— .820	82.1	—1.363	94.6	—2.384
37	pair.....	2.9	+2.811	11.2	+1.803	42.9	+ .265	77.5	—1.120
38	shoes....	1.7	+3.146	18.9	+1.307	46.4	+ .134	76.9	—1.091
39	they....	16.6	+1.438	34.3	+ .600	64.3	— .543	88.0	—1.742
40	are.....	18.3	+1.340	58.0	— .299	84.5	—1.506	95.9	—2.579
41	at.....	43.4	+ .246	72.8	— .900	75.0	—1.000	86.7	—1.649
42	all.....	36.6	+ .508	62.1	— .457	73.2	— .918	90.2	—1.918
43	pretty...	2.9	+2.811	20.1	+1.243	43.5	+ .243	78.5	—1.170
44	no.....	23.4	+1.076	71.0	— .820	85.1	—1.543	96.5	—2.686
45	man.....	50.9	— .033	74.0	— .954	79.8	—1.238	97.2	—2.155
46	ought...	1.1	+3.392	12.4	+1.713	20.8	+1.206	53.8	— .141
47	steal....	1.7	+3.146	12.4	+1.713	29.8	+ .786	67.1	— .656
48	even....	5.1	+2.425	24.3	+1.033	48.8	+ .044	67.4	— .669
49	a.....	60.6	— .399	72.8	— .900	86.3	—1.622	98.4	—3.182
50	penny...	3.4	+2.706	13.6	+1.629	35.7	+ .543	63.6	— .516

Figures 26, 27, and 28 give the scales for these words. In Figure 27 it is indicated below the line with the omission of the six words noted in the last paragraph. Above the line the words of the Preferred List are reproduced from Figure 14. Since the Easy 50-Word Test was also given to 4th-grade children it is likewise scaled for that grade omitting the same six words. (Fig. 25, 4th grade, lower line.)

We have, therefore, scales for every grade from the first half of the second grade to and including the eighth. All of these scales above the 2d grade are much richer than were those given in Section 12. There are fewer gaps in them and their range is greater. They may be used to great advantage in testing the spelling ability of children in any grade of the elementary school in which children are supposed to have any such ability. If it is not convenient to use a whole scale, certain words differing in difficulty by approximately equal amounts may be selected. Groups of words may be made each of equal difficulty as a group, or each differing from the preceding group by a fixed amount. The position of each word shows the weight which ought to be assigned to it for test purposes.

Each of these grade scales refers to the median of the grade as the zero-point. In Figure 29 is shown a scale for all grades referring, as in Figure 20, to the median of the 3d grade as the zero-point. Above the line is shown the Preferred List as in Fig 20. Below it are arranged the words of Rice's Test; and on a parallel scale the Easy 50-Word List. Caution, however, ought to be observed in accepting too literally the showing of the last two lists. Rice's Test was not given to the 3d grade, and the Easy 50-Word Test was given to the 2d grade and was not given above the 4th. They cannot, therefore, be closely compared with the Preferred List. The effect of high grades is to make the words harder, of low grades to make them easier. In the case of the Rice Test the words are probably a little—but only a little—too far to the right,— i.e., farther toward the high end than they would have been had they been used in the 3d grade—as Cornman used them. In the case of the Easy 50-Word Test the words would be a great deal too far to the left if set down as the record indicated. The six words common



← Easy 50-Word Test (Corr

the Rice List being at +685 fr

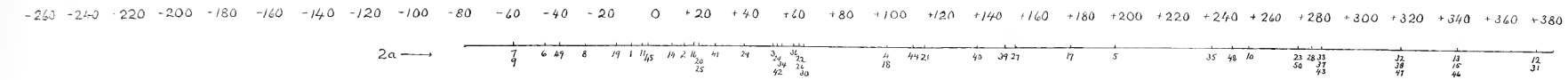


Fig 26
2a and 2b Grade Scales, Easy 50-Word List. Table XXXII

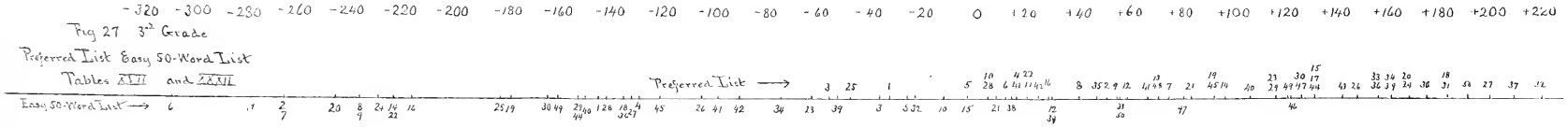


Fig 27 3rd Grade
Preferred List Easy 50-Word List
Tables XXII and XXXII

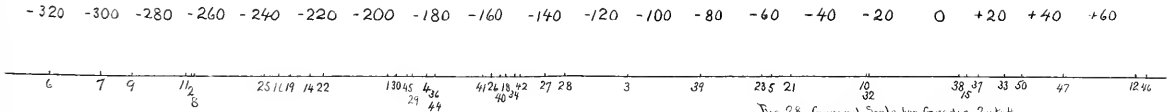


Fig 28 General Scale for Grades 2 to 4
Easy 50-Word List Table XXXII
Point of reference, 3rd Grade Median.

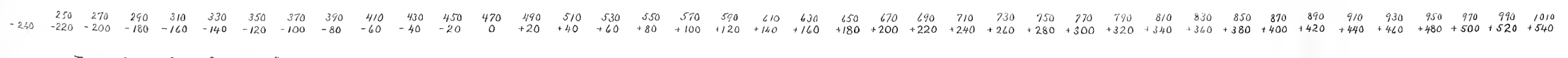
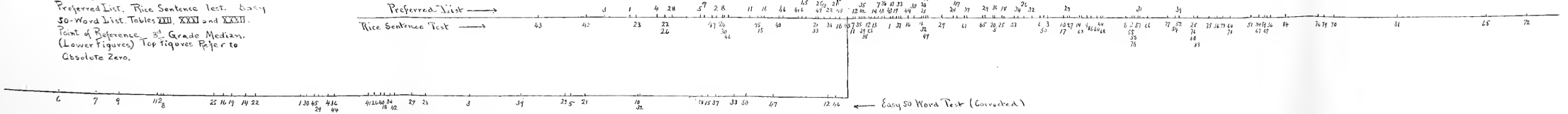


Fig 29. General Scale. Grade 2a-8th
Preferred List. Rice Sentence Test. Easy 50-Word List. Tables XXII, XXXII and XXXII.
Point of Reference, 3rd Grade Median.
(Lower Figures) Top figures Refer to Absolute Zero.



Figs. 26-29. For the various lists referred to, see Appendix II. The scale for Fig. 29 is not complete, word 82 of the Rice List being at +685 from the 3d-grade median, or +1155 from absolute zero.

Absolute Zero
(1 over Fig[illegible])

Fig. 21. 3rd degree

Fig. 20. Central Scale. Curves 20-23.

to this list and to the Preferred List enable us to suggest a correction. Their P.E. values, when the averages of the grades writing them are taken, appear as follows:

	Easy 50-word List	Preferred List	Increase
sure	+.530	+1.57	1.04
front	—.299	+1.06	1.36
touch	+.652	+1.71	1.06
only	—.089	— .57	.66
pretty	—.024	+1.31	1.33
even	—.097	+ .70	.80
		Average Increase	1.04

It appears, therefore, that in order to compare the words of the Easy 50-Word List with those of the Preferred List and to scale them together we ought to raise all the words of the former list about 1 P.E. In Figure 29, accordingly, all these words have been raised that amount.

Fig. 29 shows our most complete scale. It has decided limitations, and it is impossible—in the case of the newly added words—to suppose that it is more than an approximation. A great deal more testing than we have been able to do will have to be done before these words and others with them can be precisely fixed beyond dispute. It is not claimed that the scale we give is final. We think, however, that, supposing the two fundamental assumptions upon which it is based to be valid, it may be used in its present form with substantially accurate results; and we are confident that the general method by which it has been derived is the one by which a final scale may ultimately be secured.

The top figures in Fig. 29 refer to the absolute zero-point, taken as 470 below the 3d-grade median. It enables us to state not only the difference in difficulty between words but their relationships. We may say, for instance, that *school* (No. 27 E. 50-W. L., scales at 428) is one-half as hard as *grateful* (No. 51 R. S. T., scales at 856). We may put certain facts in equation form as follows:

in = ½ *light* = ⅓ *pigeons* = ¼ *fatiguing*
is = ½ *also* = ⅓ *occasion* = ¼ *conscientious*
the = ½ *chicken* = ⅓ *approval*
and = ½ *penny* = ⅓ *peculiar*

Many more such statements may be made. It will, we think, surprise most people to learn that *fatiguing* is only four times as hard as *he,* or that to spell *occasion* shows but three times as much ability as to spell *is.*

In fact it will, we think, be seriously questioned whether such words as *at, of, on, do,* etc., have difficulties anything like as great as is shown on our scale. It will be asked, What words can be easier than these? If a child cannot spell them does he not show zero ability? The answer is that if one or more of these very easy words were isolated and pronounced to a group to be written, those who could not spell them would indeed show no spelling ability. But these words were not isolated, they were given in a context. It is one thing for children to write the word "at" when pronounced alone or in column dictation. It is quite another to write it in the sentence: "They are not at all pretty." Some will omit it, and this fault is not confined by any means to the lowest classes. Some will connect it with the word *all,* because they habitually do so in speaking. Some little children will quite break down on the whole sentence because they can't get over the word "they." In other sentences some will substitute a word (generally of similar meaning) for the one dictated. Each of these faults scores "wrong," and none of them would be made in column dictation. It is also true that children writing sentences more often write illegibly than they do when writing a few words in columns; and this is particularly true with young children.

It will therefore be clear that the decision as to how hard a word is, depends on how you use it in testing and when you call it "wrong." To verify the placings of the words given in this study one ought to use the same test material and the same method of scoring. In particular, column dictation will not do at all.

§ 19. *Derived Forms of Distribution*

The foregoing treatment of the measurement of spelling ability has, as has been indicated frequently, proceeded upon the assumption that the distribution of ability is in all grades normal. Such an assumption has always been made in the investigation of school abilities by persons whose knowledge of

the theory of statistics has enabled them to do so. In Section 10 I have said: "There seems no good ground for assuming that the distribution of spelling ability in any grade is not according to the normal curve *or according to a curve which resembles it closely.*" By this alternative is suggested the possible applicability of certain curves not of normal form but resembling the normal form. Our problem will now be to derive and apply to some of our material such modifications of the type form of distribution as our present knowledge of grade conditions permits.

In order that the frequency of measurements within a group may be distributed according to the Probability Integral it is necessary that the group be in no way selected on the basis of the characteristic that is measured. It must be a random sampling from a "total population." If the frequency distributions of statures for adult males born in the City of New York may be expected to approximate the symmetrical type, the distributions of statures for adult males on the police force of New York City would not do so. Their curve will be of the "moderately asymmetrical type" being cut off at the low end because extremely short men are at a disadvantage in the group supposed to be measured. In other words, there is a selection on the basis of stature. The group "adult males on the police force of New York City" is not a random sampling of the total population "adult males of New York City."

The question then is: To what extent does the membership of each grade of the elementary school fail of being a chance selection from a total population? We may fairly assume, in the first place, that the pupils of the first and second grades are unselected. Practically all children attend school and none drop out in these grades. From the 3d grade on, however, each successive grade constitutes a group which is less and less a random sampling. Many influences are at work to eliminate a greater and greater number of individuals. Probably the most important of them is the inability of children to progress—i.e., lack of ability in the lines of work now required by the schools.

The extent to which elimination takes place in the grades has been the subject of study by a number of investigators. The first of these was Thorndike ('07). He draws conclusions from

conditions in 23 cities as they were about 1900. He estimates that out of 100 entering pupils, 97 remain till grade 3, 90 till grade 4, 81 till grade 5, 68 till grade 6, 54 till grade 7, and 40 till the last grammar grade (8th or 9th). Ayres ('09) sharply criticised these figures, stating that they were too small. He contended, particularly, that there was no dropping out before the 6th year—a conclusion which common observation and later investigation unite to disprove. Employment certificates are granted in great numbers to 5th-grade children. Mr. Ayres' figures for retention are as follows: Grades 1-5, 100 (i.e., no elimination); grade 6, 90; grade 7, 71; grade 8, 51.

Thorndike, using later and better reports, subsequently derived figures a little higher than his former ones, but substantially in agreement with them (Thorndike, '10). They were no higher probably than 5 or 6 years of agitation would have led one to expect.

Another important study of this question was made by Strayer ('11), the material being used from 318 cities. His conclusions tend to group with Thorndike's rather than with those of Ayres. Owing to the large number of cities whose returns were used, the uniform method of taking the census, and the recency of the conditions studied, this investigation is highly important. No single figures are given for retention in general, though they are easily found. Using the largest age group as the number of entering children, he gives the following as the median per cents in each grade.

	Cities of Over 25,000		Cities of Less than 25,000	
	Boys	Girls	Boys	Girls
3d year	115	110	110	105
4th "	110	110	105	100
5th "	100	95	95	95
6th "	85	85	80	85
7th "	65	75	70	70
8th "	50	60	50	60

Since in this study we group boys and girls together and consider general conditions, the average of these percentages will give figures for retention for each grade (subject to deduction for

repeaters) as follows: 3d grade, 110; 4th grade, 106; 5th grade, 96; 6th grade, 84; 7th grade, 70; 8th grade, 55. If, as Dr. Strayer says, a fair estimate of the number of repeaters in the 6th, 7th, and 8th grades would be 12%, 10%, and 8% of the pupils in each grade (p. 136), it is likely that the progression (8, 10, 12) may be carried back to the 5th, 4th and 3d grades without great violence to the facts. We estimate therefore that the number of repeaters in the 3d, 4th, and 5th grades is 18%, 16%, and 14% of the pupils in each grade. Making these deductions from the above percentages, we have for the retention: 3d grade, 92; 4th grade, 90; 5th grade, 82; 6th grade, 72; 7th grade, 60; 8th grade, 47.

Weighing as best we can the results of these four studies, we have made the best estimate we can for the probable amount of retention at present in the grades. For reasons that will appear later we have expressed this estimate in numbers per 10,000 instead of per 100.

Table XXXIII and Fig. 30 show the percentages we have adopted compared with those of Thorndike, Ayres, and Strayer (as derived). Fig. 30 gives only the earlier of Thorndike's percentages.

TABLE XXXIII

PERCENTAGES OF RETENTION. GRADES 3 TO 8

	3d	4th	5th	6th	7th	8th
Thorndike '07	97	90	81	68	54	40
Ayres '09	100	100	100	90	71	51
Thorndike '10		91	81.5	70.9	56	41.2
Strayer '11 (derived)	92	90	82	72	60	47
Adopted	97.25	95.46	88.40	70.87	57.44	48.21

Such an amount of retention for each grade having been adopted, the next question to consider is: What part of a normal distribution is thus eliminated? Obviously not all the poorest in ability drop out. Our results for spelling show that some very poor spellers are retained even in the highest grades. Yet the greatest elimination will no doubt be among those of lowest ability and will be progressively less among children of greater ability. How much this amounts to for successive incre-

ments of ability we do not positively know. We are again forced to make as reasonable an estimate as we can, and this time without the help of any investigations.

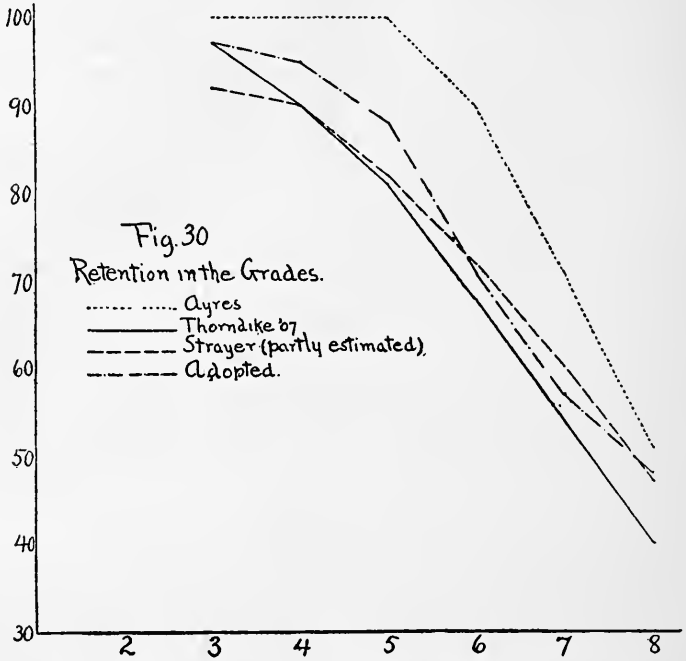


FIG. 30. The horizontal scale is for the grades of the elementary school; the vertical scale is for percentage of pupils retained.

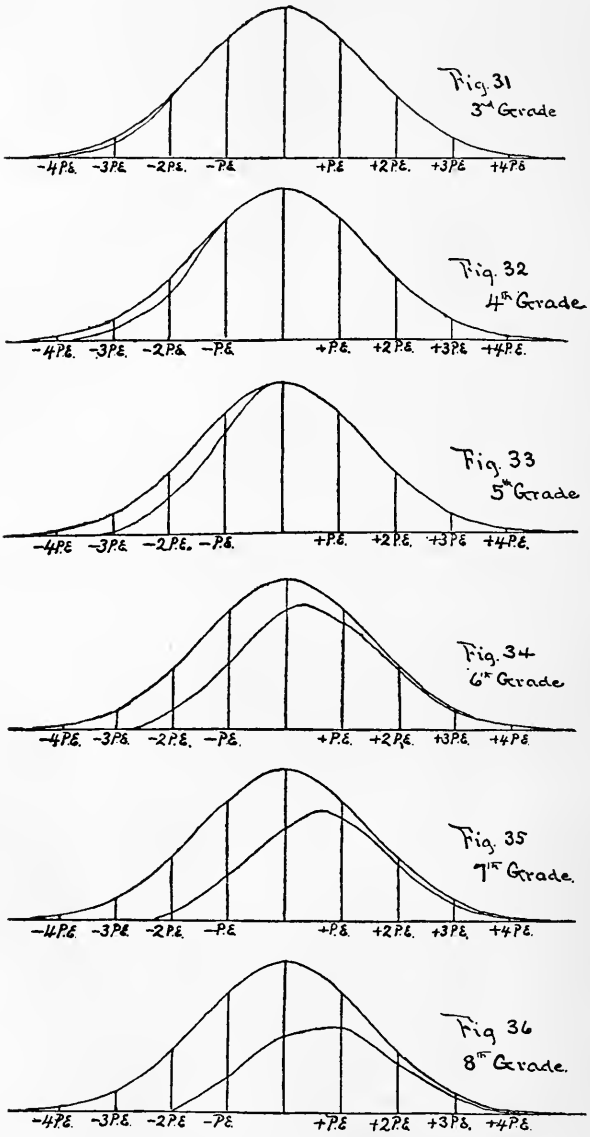
We have estimated (Table XXXIII) that 9725 out of 10,000 entering children are retained in the 3d grade; and we judge further that all below — 4 P.E. have dropped out, that at — 3 P.E. 40% have dropped out, and that at — 2 P.E. none at all. In the 4th grade, retention is put at 9546 in 10,000 (Table XXXIII) and total elimination is estimated to operate up to— 3.7 P.E. From this point to — 2.7 P.E. the force of elimination is supposed gradually to diminish until 40% drop out. At — 1.7 P.E. only 10% are estimated as lost to the grade. At — 1.2 P.E. the forces tending to eliminate children are supposed to have been completely counteracted by the opposing forces tending to retain them. In the 5th grade (8840 out of 10,000 retained) the elimination is

assumed to be total up to — 3.2 P.E. and partial as high as — 0.2 P.E. In the 6th grade (7087 out of 10,000 retained) the corresponding points are — 2.7 P.E. and + 3.3 P.E.; in the 7th (5744 retained) they are — 2.3 P.E. and + 3.7 P.E.; and in the 8th (4821 retained), — 2 P.E. and + 6 P.E. It is, therefore, supposed that in the last three years of the elementary school there is some elimination even among the most capable children. This plan of elimination and retention may easily be attacked as artificial and it may very likely be shown to need considerable modification when later and better knowledge is available on this difficult subject. Meanwhile, however, some assumption was necessary in order to construct any sort of frequency tables which should illustrate the method of constructing a scale when account is taken of the selective influence of the grades. We can only state that we have keenly appreciated the importance of distributing the amount of elimination where it most probably occurs, that we have been at no small pains to find out where to distribute it, and that we have estimated as wisely as we could. Table XXXIV gives the entire plan of elimination and retention for each grade.

TABLE XXXIV

PLAN OF ELIMINATION AND RETENTION FOR EACH GRADE

Grade, etc.	X / P.E.	Per cent eliminated	Per cent retained	Grade, etc.	X / P.E.	Per cent eliminated	Per cent retained
3d Grade.	—4	100	0	7th Grade	—2.3	100	0
N=9725 .	—3	40	60	N=5744 .	—1.3	70	30
	—2	0	100		—0.3	50	50
					+0.7	20	80
4th Grade	—3.7	100	0		+1.7	10	90
N=9546 .	—2.7	40	60		+2.7	5	95
	—1.7	10	90		+3.7	0	100
	—1.2	0	100				
				8th Grade	—2	100	0
5th Grade	—3.2	100	0	N=4821 .	—1	70	30
N=8840 .	—2.2	50	50		0	50	50
	—1.2	20	80		+1	30	70
	—0.2	0	100		+2	20	80
					+3	10	90
6th Grade	—2.7	100	0		+4	6	94
N=7087 .	—1.7	60	40		+5	2	98
	—0.7	35	65		+6	0	100
	+0.3	15	85				
	+1.3	10	90				
	+2.3	5	95				
	+3.3	0	100				



FIGS. 31-36. The estimated amount and distribution of elimination and retention. See Table XXXIII.

The next step was to apply the data of Table XXXIV to the normal distribution and to derive therefrom for each grade a modified distribution which should take account of the amount and range of elimination as estimated. In order that the validity of our method may be open to inspection, we shall illustrate for the 6th grade the manner in which these modified distributions were derived.

We have adopted certain percentages of retention for designated amounts of general ability (Table XXXIV, 6th grade), and these percentages must not only stand the test of reasonableness in themselves, but they must also when applied to a normal table of frequency (the sum of whose cases is, say, 1000), reduce the number of cases to an amount which represents a reasonable percentage of retention for the 6th grade (say, 70 or 71). That is, the derived table must show approximately 700 cases out of 1000, or 7000 out of 10,000. We shall see later to what extent this turns out to be true.

Adapting the normal table of frequency (Table XIV, page 35) so as to include 1000 cases instead of 10,000 and taking intervals of 0.1 P.E., we have columns 1 and 2 of Table XXXV. In column 3 we increase the percentages of retention from 0 at — 2.7 P.E. to 40 at — 1.7 P.E. by increments of 4 for each of the ten steps; then by increments of 2.5 until 65 is reached at — 0.7 P.E.; then by increments of 2 to 85 at + 0.3 P.E.; and so on as required by Table XXXIV, col. 4. Taking these percentages of the frequencies in column 2 gives the derived frequencies of column 4. The sum of the entries in this column being 708.7, the plan gives an amount of elimination which is reasonable for the 6th grade. (See Table XXXIII.)

The amount and distribution of elimination and retention are shown by diagram for each grade in Figs. 31 to 36. Fig. 34 in particular shows these facts for the 6th grade, and is the graphic representation of the series of frequencies in column 4 of Table XXXV. Fig. 31 shows the same facts for the 3d grade, Fig. 32 for the 4th grade, etc. The progressive increase in elimination and the extension of it to higher and higher parts of the normal curve are the facts to be noticed.

But we have not in column 4 of Table XXXV, a frequency table for the 6th grade in the most useful form. The area of its

TABLE XXXV

SIXTH GRADE. DERIVATION OF MODIFIED TABLE OF FREQUENCY

Below Normal Median					Above Normal Median				
X / P.E.	Normal Frequencies	Percentages of Retention	Derived Frequencies	Same on basis of 10,000 cases	X / P.E.	Normal Frequencies	Percentages of Retention	Derived Frequencies	Same on basis of 10,000 cases
0—.1	27	79	21.3	301	0—.1	27	81	21.9	309
.2	27	77	20.8	293	.2	27	83	22.4	312
.3	26	75	19.5	275	.3	26	*85*	22.1	316
.4	26	73	19.0	268	.4	26	85.5	22.2	316
.5	26	71	18.5	261	.5	26	86	22.4	313
.6	25	69	17.3	244	.6	25	86.5	21.6	308
.7	25	67	16.8	237	.7	25	87	21.8	305
.8	23	*65*	15.0	212	.8	23	87.5	20.1	284
.9	23	62.5	14.4	203	.9	23	88	20.2	285
1.0	22	60	13.2	186	1.0	22	88.5	19.5	275
1.1	21	57.5	12.1	171	1.1	21	89	18.7	264
1.2	20	55	11.0	155	1.2	20	89.5	17.9	252
1.3	19	52.5	10.0	141	1.3	19	*90*	17.1	241
1.4	18	50	9.0	127	1.4	18	90.5	16.3	230
1.5	16	47.5	7.6	107	1.5	16	91	14.6	206
1.6	16	45	7.2	102	1.6	16	91.5	14.6	206
1.7	14	42.5	6.0	85	1.7	14	92	12.9	182
1.8	14	*40*	5.6	79	1.8	14	92.5	13.0	183
1.9	12	36	4.3	61	1.9	12	93	11.2	158
2.0	11	32	3.5	50	2.0	11	93.5	10.3	145
2.1	11	28	3.1	44	2.1	11	94	10.3	145
2.2	9	24	2.2	31	2.2	9	94.5	8.5	121
2.3	9	20	1.8	26	2.3	9	*95*	8.6	120
2.4	7	16	1.1	16	2.4	7	95.5	6.7	95
2.5	7	12	.8	11	2.5	7	96	6.7	95
2.6	6	8	.5	7	2.6	6	96.5	5.8	82
2.7	6	4	.2	3	2.7	6	97	5.8	82
2.8	5	*0*	0	0	2.8	5	97.5	4.9	69
					2.9	4	98	3.9	55
					3.0	4	98.5	3.9	55
					3.1	3	99	3.0	42
					3.2	3	99.5	3.0	42
					3.3	2	*100*	2.0	28
					3.4	2	100	2.0	28
					3.5	2	100	2.0	28
					3.6	2	100	2.0	28
					3.7	1	100	1.0	14
					3.8	1	100	1.0	14
					3.9	1	100	1.0	14
					4.0	1	100	1.0	14
					etc. to 6.0	etc. to .02	etc.	etc. to .02	etc. to .28
Total No. cases.....						1000		708.7	9999.72

curve is no longer 1000, but only 708.7. In order to express the several frequencies in the form of per cents, we shall have to divide each of them (column 4) by their total (708.7). Expressing these quotients on the basis of 10,000 instead of 1000, we have column 5. These are the numbers in the columns 3 and 5 of Table XXXIX (p. 96); and when their sums are taken beginning at 0 they constitute the Modified Table of Frequency for the 6th grade (Table XXXIX).

TABLE XXXVI

MODIFIED TABLE OF FREQUENCY, 3D GRADE. MEDIAN = +0.051 P.E.

Plan of elimination: —4 P.E., 100%; —3 P.E., 40%; —2 P.E., 0%. Total area of the surface of frequency taken as 10,000. See Fig. 37.

X	Low		High		X	Low		High		X	Low		High	
P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ
		278		278			109		113					5.1
.1	278		278		2.1	4334		4338		4.1			5116.1	
		278		278			85		93					5.1
.2	556		556		2.2	4419		4431		4.2			5121.2	
		267		267			81		93					5.1
.3	823		823		2.3	4500		4524		4.3			5126.3	
		267		267			60		72					4.1
.4	1090		1090		2.4	4560		4596		4.4			5130.4	
		267		267			58		72					2.1
.5	1357		1357		2.5	4618		4668		4.5			5132.5	
		257		257			47		62					2.1
.6	1614		1614		2.6	4665		4730		4.6			5134.6	
		257		257			44		62					1.0
.7	1871		1871		2.7	4709		4792		4.7			5135.6	
		236		236			35		51					1.0
.8	2107		2107		2.8	4744		4843		4.8			5136.6	
		236		236			26		41					1.0
.9	2343		2343		2.9	4770		4884		4.9			5137.6	
		226		226			25		41					1.0
1.0	2569		2569		3.0	4795		4925		5.0			5138.6	
		216		216			17		31					.51
1.1	2785		2785		3.1	4812		4956		5.1			5139.11	
		206		206			15		31					.51
1.2	2991		2991		3.2	4827		4987		5.2			5139.62	
		195		195			9		21					.31
1.3	3186		3186		3.3	4836		5008		5.3			5139.93	
		185		185			7.4		21					.31
1.4	3371		3371		3.4	4843.4		5029		5.4			5140.24	
		165		165			6.2		21					.31
1.5	3536		3536		3.5	4849.6		5050		5.5			5140.55	
		165		165			4.9		21					.31
1.6	3701		3701		3.6	4854.5		5071		5.6			5140.86	
		144		144			1.9		10					.21
1.7	3845		3845		3.7	4856.4		5081		5.7			5141.07	
		144		144			1.2		10					.21
1.8	3989		3989		3.8	4857.6		5091		5.8			5141.28	
		123		123			0.6		10					.21
1.9	4112		4112		3.9	4858.2		5101		5.9			5141.49	
		113		113					10					.21
2.0	4225		4225		4.0			5111		6.0			5141.70	

TABLE XXXVII

MODIFIED TABLE OF FREQUENCY, 4TH GRADE. MEDIAN=+0.087 P.E.

Plan of elimination: —3.7 P.E., 100%; —2.7 P.E., 40%; —1.7 P.E., 10%; —1.2 P.E., 0%. Total area of the surface of frequency taken as 10,000. See Fig. 38.

X	Low		High		X	Low		High		X	Low		High	
P.E.	%	△	%	△	P.E.	%	△	%	△	P.E.	%	△	%	△
		283		283			93		115					5.2
.1	283		283		2.1	4316		4421		4.1			5210.2	
		283		283			74		94					5.2
.2	566		566		2.2	4390		4515		4.2			5215.4	
		272		272			71		94					5.2
.3	838		838		2.3	4461		4609		4.3			5220.6	
		272		272			53		73					4.2
.4	1110		1110		2.4	4514		4682		4.4			5224.8	
		271		271			51		73					2.1
.5	1381		1381		2.5	4565		4755		4.5			5226.9	
		262		262			41		63					2.1
.6	1643		1643		2.6	4606		4818		4.6			5229	
		262		262			40		63					1.05
.7	1905		1905		2.7	4646		4881		4.7			5230.05	
		241		241			31		52					1.05
.8	2146		2146		2.8	4677		4933		4.8			5231.10	
		241		241			23		42					1.05
.9	2387		2387		2.9	4700		4975		4.9			5232.15	
		230		230			20		42					1.05
1.0	2617		2617		3.0	4720		5017		5.0			5233.20	
		220		220			13		31					.52
1.1	2837		2837		3.1	4733		5048		5.1			5233.72	
		210		210			11		31					.52
1.2	3047		3047		3.2	4744		5079		5.2			5234.24	
		199		199			6		21					.31
1.3	3246		3246		3.3	4750		5100		5.3			5234.55	
		185		189			5		21					.31
1.4	3431		3435		3.4	4755		5121		5.4			5234.86	
		161		168			4		21					.31
1.5	3592		3603		3.5	4759		5142		5.5			5235.17	
		158		168			3		21					.31
1.6	3750		3771		3.6	4762		5163		5.6			5235.48	
		135		147			1		10.5					.21
1.7	3885		3918		3.7	4763		5173.5		5.7			5235.69	
		132		147					10.5					.21
1.8	4017		4065		3.8			5184		5.8			5235.90	
		109		126					10.5					.21
1.9	4126		4191		3.9			5194.5		5.9			5236.11	
		97		115					10.5					.21
2.0	4223		4306		4.0			5205		6.0			5236.32	

TABLE XXXVIII

MODIFIED TABLE OF FREQUENCY, 5TH GRADE. MEDIAN=+0.215 P.E.

Plan of elimination: —3.2 P.E., 100%; —2.2 P.E., 50%; —1.2 P. E., 20%; —0.2 P.E., 0%. Total area of the surface of frequency taken as 10,000. See Fig. 39.

X	Low		High		X	Low		High		X	Low		High	
P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ
		305		305			70		124					6
.1	305		305		2.1	4093		4772		4.1			5627	
		305		305			54		102					6
.2	610		610		2.2	4147		4874		4.2			5633	
		294		294			51		102					6
.3	904		904		2.3	4198		4976		4.3			5639	
		288		294			36		79					5
.4	1192		1198		2.4	4234		5055		4.4			5644	
		282		294			32		79					2
.5	1474		1492		2.5	4266		5134		4.5			5646	
		266		283			24		68					2
.6	1740		1775		2.6	4290		5202		4.6			5648	
		260		283			20		68					1
.7	2000		2058		2.7	4310		5270		4.7			5649	
		234		260			14		57					1
.8	2234		2318		2.8	4324		5327		4.8			5650	
		229		260			9		45					1
.9	2463		2578		2.9	4333		5372		4.9			5651	
		214		249			7		45					1
1.0	2677		2827		3.0	4340		5417		5.0			5652	
		200		238			3		34					.6
1.1	2877		3065		3.1	4343		5451		5.1			5652.6	
		186		226			2		34					.5
1.2	3063		3291		3.2	4345		5485		5.2			5653.1	
		172		215					23					.3
1.3	3235		3506		3.3			5508		5.3			5653.4	
		157		204					23					.3
1.4	3392		3710		3.4			5531		5.4			5653.7	
		134		181					23					.3
1.5	3526		3891		3.5			5534		5.5			5654.0	
		129		181					23					.3
1.6	3655		4072		3.6			5577		5.6			5654.3	
		108		158					11					.2
1.7	3763		4230		3.7			5588		5.7			5654.5	
		103		158					11					.2
1.8	3866		4388		3.8			5599		5.8			5654.7	
		84		136					11					.2
1.9	3950		4524		3.9			5610		5.9			5654.9	
		73		124					11					.2
2.0	4023		4648		4.0			5621		6.0			5655.1	

TABLE XXXIX

MODIFIED TABLE OF FREQUENCY, 6TH GRADE. MEDIAN = +0.418P.E.

Plan of elimination: —2.7 P.E., 100%; —1.7 P.E., 60%; —0.7 P.E., 35%; +0.3 P.E., 15%; +1.3 P.E., 10%; +2.3 P.E., 5%; +3.3 P.E., 0%. Total area of surface of frequency taken as 10,000. See Fig. 40.

X	Low		High		X	Low		High		X	Low		High	
P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ
		301		309			44		145					7
.1	301		309		2.1	3602		5235		4.1			6268	
		293		312			31		121					7
.2	594		621		2.2	3633		5356		4.2			6275	
		275		316			26		120					7
.3	869		937		2.3	3659		5476		4.3			6282	
		268		316			16		95					6
.4	1137		1253		2.4	3675		5571		4.4			6288	
		261		313			11		95					3
.5	1398		1566		2.5	3686		5666		4.5			6291	
		244		308			7		82					3
.6	1642		1874		2.6	3693		5748		4.6			6294	
		237		305			3		82					1.4
.7	1879		2179		2.7	3696		5830		4.7			6295.4	
		212		284					69					1.4
.8	2091		2463		2.8			5899		4.8			6296.8	
		203		285					55					1.4
.9	2294		2748		2.9			5954		4.9			6298.2	
		186		275					55					1.4
1.0	2480		3023		3.0			6009		5.0			6299.6	
		171		264					42					.7
1.1	2651		3287		3.1			6051		5.1			6300.3	
		155		252					42					.7
1.2	2806		3539		3.2			6093		5.2			6301	
		141		241					28					.4
1.3	2947		3780		3.3			6121		5.3			6301.4	
		127		230					28					.4
1.4	3074		4010		3.4			6149		5.4			6301.8	
		107		206					28					.4
1.5	3181		4216		3.5			6177		5.5			6302.2	
		102		206					28					.4
1.6	3283		4422		3.6			6205		5.6			6302.6	
		85		182					14					.28
1.7	3368		4604		3.7			6219		5.7			6302.88	
		79		183					14					.28
1.8	3447		4787		3.8			6233		5.8			6303.16	
		61		158					14					.28
1.9	3508		4945		3.9			6247		5.9			6303.44	
		50		145					14					.28
2.0	3558		5090		4.0			6261		6.0			6303.72	

TABLE XL

MODIFIED TABLE OF FREQUENCY, 7TH GRADE. MEDIAN=+0.669 P.E.

Plan of elimination: —2.3 P.E., 100%; —1.3 P.E., 70%; —0.3 P.E., 50%; +0.7 P.E., 20%; +1.7 P.E., 10%; +2.7 P.E., 5%; +3.7 P.E., 0%. Total area of frequency surface taken as 10,000. See Fig. 41.

X	Low		High		X	Low		High		X	Low		High	
P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ
		277		290			17		176					9
.1	277		290		2.1	2844		5846		4.1			7096	
		263		306			9		145					9
.2	540		596		2.2	2853		5991		4.2			7105	
		240		308			5		145					9
.3	780		904		2.3	2858		6136		4.3			7114	
		226		322					114					7
.4	1006		1226		2.4			6250		4.4			7121	
		217		334					114					3
.5	1223		1560		2.5			6364		4.5			7124	
		200		336					99					3
.6	1423		1896		2.6			6463		4.6			7127	
		191		348					99					2
.7	1614		2244		2.7			6562		4.7			7129	
		169		326					83					2
.8	1783		2570		2.8			6645		4.8			7131	
		160		326					66					2
.9	1943		2896		2.9			6711		4.9			7133	
		146		318					68					2
1.0	2089		3214		3.0			6779		5.0			7135	
		132		306					50					.9
1.1	2221		3520		3.1			6829		5.1			7135.9	
		118		296					50					.9
1.2	2339		3816		3.2			6879		5.2			7136.8	
		106		283					35					.5
1.3	2445		4099		3.3			6914		5.3			7137.3	
		94		273					35					.5
1.4	2539		4372		3.4			6949		5.4			7137.8	
		75		246					35					.5
1.5	2614		4618		3.5			6984		5.5			7138.3	
		66		246					35					.5
1.6	2680		4864		3.6			7019		5.6			7138.8	
		50		220					17					.35
1.7	2730		5084		3.7			7036		5.7			7139.15	
		43		220					17					.35
1.8	2773		5304		3.8			7053		5.8			7139.50	
		31		190					17					.35
1.9	2804		5494		3.9			7070		5.9			7139.85	
		23		176					17					.3
2.0	2827		5670		4.0			7087		6.0			7140.15	

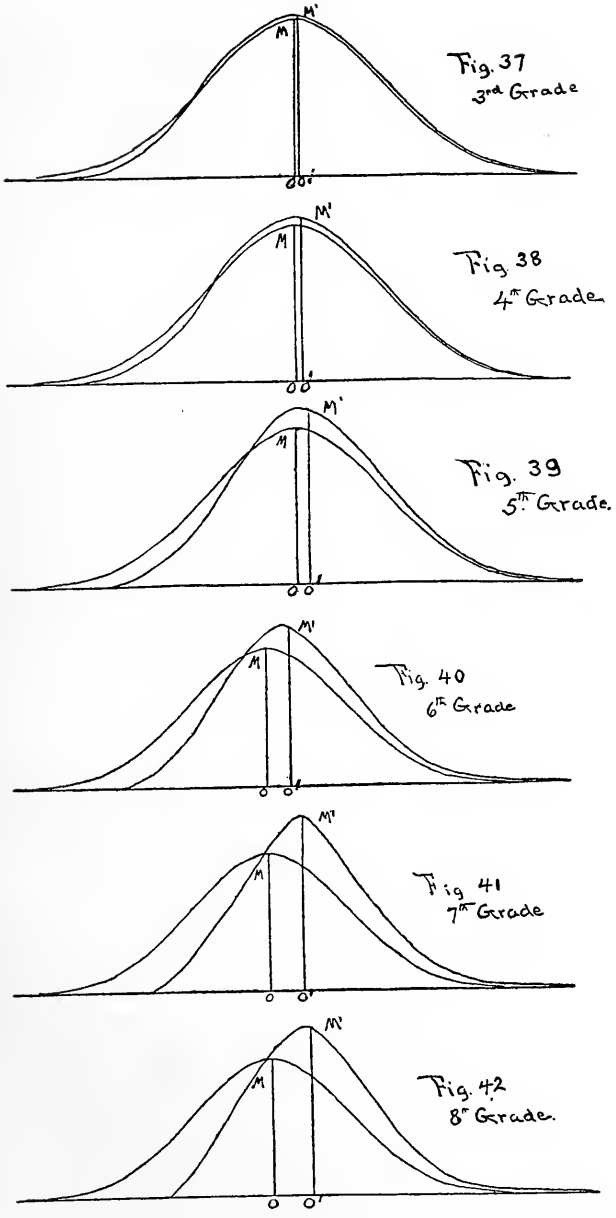
TABLE XLI

MODIFIED TABLE OF FREQUENCY, 8TH GRADE. MEDIAN=+0.746 P.E.

Plan of elimination: —2 P.E., 100%; —1 P.E., 70%; 0 P.E., 50%; +1 P.E., 30%; +2 P.E., 20%; +3 P.E., 10%; +4 P.E., 6%; +5 P.E., 2%; +6 P.E., 0%. Total area of frequency surface taken as 10,000. See Fig. 42.

X	Low		High		X	Low		High		X	Low		High	
P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ	P.E.	%	Δ	%	Δ
		280		290					184					10
.1	280		290		2.1			5834		4.1			7195	
		270		303					165					10
.2	550		593		2.2			5999		4.2			7205	
		249		303					145					10
.3	799		896		2.3			6144		4.3			7215	
		237		313					130					8
.4	1036		1209		2.4			6274		4.4			7223	
		226		323					116					4
.5	1262		1532		2.5			6390		4.5			7227	
		207		323					108					4
.6	1469		1855		2.6			6498		4.6			7231	
		197		332					108					2
.7	1666		2187		2.7			6606		4.7			7233	
		172		320					91					2
.8	1838		2507		2.8			6697		4.8			7235	
		162		319					75					2
.9	2000		2826		2.9			6772		4.9			7237	
		145		319					75					2
1.0	2145		3145		3.0			6847		5.0			7239	
		131		309					56					1
1.1	2276		3454		3.1			6903		5.1			7240	
		112		299					56					1
1.2	2388		3753		3.2			6959		5.2			7241	
		95		288					38					.6
1.3	2483		4041		3.3			6997		5.3			7241.6	
		79		276					38					.6
1.4	2562		4317		3.4			7035		5.4			7242.2	
		60		260					37					.6
1.5	2622		4577		3.5			7072		5.5			7242.8	
		50		242					37					.6
1.6	2672		4819		3.6			7109		5.6			7243.4	
		35		235					19					.4
1.7	2707		5054		3.7			7128		5.7			7243.8	
		27		215					19					.4
1.8	2734		5269		3.8			7147		5.8			7244.2	
		14		197					19					.4
1.9	2748		5466		3.9			7166		5.9			7244.6	
		6		184					19					.4
2.0	2754		5650		4.0			7185		6.0			7245	

In this manner each of the Modified Tables of Frequency was made up. They are given in tables XXXVI to XLI. They are intended to take the place, each for the grade to which it applies, of the Table of Frequency for the normal distribution. Since they are asymmetrical, the lower and upper parts have to be given separately. For the same reason there is no P.E., the use of the Probable Error as a unit of amount being properly confined to normal curves only (Yule, '11, p. 147). The quartile deviation $\frac{(Q_3 - Q_1)}{2}$ might be used instead, but its



FIGS. 37–42. Derived Forms of Distribution. Grades 3 to 8.

value differs for each of the six tables. In order therefore to employ a unit which should be the same for all, including the normal distribution, we have retained the P.E. of the Probability Integral. It is now no longer a function of the modified distributions, but a mere unit of length. Likewise in order to have a common point of reference the median of the normal distribution has been retained, the terms "low" and "high" in the tables referring to parts below or above that point. The real median of each modified distribution, however, is given, being expressed as a deviation from the old median.

Figs. 37 to 42 are to be considered in connection with tables XXXVI to XLI, of which they are the graphic expressions (the curves being "smoothed," to represent an indefinite number of cases). They are also to be considered in connection with Figs. 31 to 36, from the "retention" parts of which they are derived by making the areas 10,000. In Figs. 37 to 42, the curve extending farther to the left is in each case the normal curve and OM is its median vertical. O^1M^1 is the median vertical of the modified surface and OO^1 is the distance between medians. The values of these are as follows: 3d grade, 0.051 P.E.; 4th grade, 0.087 P.E.; 5th grade, 0.215 P.E.; 6th grade, 0.418 P.E.; 7th grade, 0.669 P.E.; 8th grade, 0.746 P.E.

Is it worth while to use these tables instead of the normal one? Will the same material when analyzed by the skew and normal distributions yield differences that are important? With the purpose of throwing some light on this question we have used the modified tables to interpret the results of testing with our Preferred List, and the rest of the present section will be devoted to this matter. The differences will not in many cases be found to be large. This is, of course, particularly true when the early grades are concerned, the curves being for those grades almost normal. It may be remarked, however, that the applicability of these tables does not rest upon the results here shown. It is general ability rather than spelling ability that tends strongly to keep children in school. Spelling ability does not correlate as highly with general ability as do the abilities in most other school subjects. It is quite likely, therefore, that

the use of these tables for the statistical treatment of other subjects may be more satisfactory than it is for spelling. They are given here primarily to illustrate the method.

TABLE XLII

NUMBER AND PER CENT OF PUPILS IN EACH GRADE WHOSE ABILITY EQUALLED OR EXCEEDED THAT OF THE MEDIAN PUPIL IN EVERY OTHER GRADE WITH THE P.E. VALUES CORRESPONDING TO EACH PER CENT. SELECTED LIST. MODIFIED DISTRIBUTIONS. COMPARE WITH TABLE XV (P. 36)

		3d Grade	4th Grade	5th Grade	6th Grade	7th Grade	8th Grade
3d grade...	No.		76	27	9	3	0
N=445	%		17.1	6.1	2.0	0.7	0
	P.E.		1.3858	2.2736	3.0029	3.607	?
4th grade..	No.	378		146	52	27	9
N=467	%	80.9		31.3	11.1	5.8	1.9
	P.E.	1.1949		.6965	1.7616	2.2778	3.0076
5th grade..	No.	478	370		142	73	30
N=515	%	92.8	71.8		27.6	14.2	5.8
	P.E.	1.7916	.7341		.8136	1.5172	2.2104
6th grade..	No.	414	384	338		142	57
N=418	%	99.0	91.9	80.1		34.0	13.6
	P.E.	2.5043	1.6747	1.0451		.5386	1.4812
7th grade..	No.	363	354	328	256		99
N=365	%	99.5	96.4	89.9	70.1		27.1
	P.E.	2.6038	2.0360	1.5096	.6586		.7564
8th grade..	No.	227	276	269	241	200	
N=277	%	100	99.6	97.1	87.0	72.2	
	P.E.	?	2.4719	2.0197	1.4424	.635	

In Section 11 we located the grade medians assuming normal distribution. In Tables XLII and XLIII the same data have been subjected to analysis using the modified distributions. These tables are to be compared with Tables XV (page 36) and XVI (page 39). The median intervals are considerably less than they were found to be by using the normal distribution. Note the comparisons in Table XLIV (page 103).

On the average, the intervals by the present method are less than the same intervals found by using the normal distribution by 0.1247 P.E., or about half a step in the 10-point scale (Table XX, page 52). Since this occurs five times the entire range

TABLE XLIII

DIRECT AND DERIVED VALUES OF MEDIAN DISTANCES. MODIFIED DISTRIBUTIONS. SELECTED LIST

	M_{3-4}	M_{4-5}	M_{5-6}	M_{6-7}	M_{7-8}
	1.3858 (direct)	.8878 (M_{3-5}—M_{3-4})	.7293 (M_{3-6}—M_{3-5})	.6041 (M_{3-7}—M_{3-6})	? (M_{3-8}—M_{3-7})
	1.5771 (M_{3-5}—M_{4-5})	.6965 (direct)	1.0651 (M_{4-6}—M_{4-5})	.5162 (M_{4-7}—M_{4-6})	.7298 (M_{4-8}—M_{4-7})
	1.2413 (M_{3-6}—M_{4-6})	.9478 (M_{4-6}—M_{5-6})	.8136 (direct)	.7036 (M_{5-7}—M_{5-6})	.6932 (M_{5-8}—M_{5-7})
	1.3292 (M_{3-7}—M_{4-7})	.7606 (M_{4-7}—M_{5-7})	.9786 (M_{5-7}—M_{6-7})	.5386 (direct)	.9426 (M_{6-8}—M_{6-7})
	? (M_{3-8}—M_{4-8})	.7972 (M_{4-8}—M_{5-8})	.7292 (M_{5-8}—M_{6-8})	.7248 (M_{6-8}—M_{7-8})	.7564 (direct)
	1.1949 (direct)	.5967 (M_{5-3}—M_{4-3})	.7127 (M_{6-3}—M_{5-3})	.0995 (M_{7-3}—$M_{\ -3}$)	? (M_{8-3}—M_{7-3})
	1.0575 (M_{5-3}—M_{5-4})	.7341 (direct)	.9406 (M_{6-4}—M_{5-4})	.3613 (M_{7-4}—M_{6-4})	.4359 (M_{8-4}—M_{7-4})
	.8296 (M_{6-3}—M_{6-4})	.9406 (M_{6-4}—M_{6-5})	1.0451 (direct)	.4645 (M_{7-5}—M_{6-5})	.5101 (M_{8-5}—M_{7-5})
	.5678 (M_{7-3}—M_{7-4})	.5264 (M_{7-4}—M_{7-5})	.8510 (M_{7-5}—M_{7-6})	.5386 (direct)	.7838 (M_{8-6}—M_{7-6})
	? (M_{8-3}—M_{8-4})	.4522 (M_{8-4}—M_{8-5})	.5773 (M_{8-5}—M_{8-6})	.8074 (M_{8-6}—M_{8-7})	.6350 (direct)
Average..	1.1479	.7340	.8443	.5359	.6858
Weighted Average	1.2008	.7483	.8685	.5606	.7065

from M_3 to M_8 is contracted by 0.7235 P.E., an amount which is more than some grade intervals (M_{6-7} by normal dis., M_{6-7} and M_{7-8} by modified dis.). This is an important difference.

In the matter of scaling the words, there is, as might be supposed, very little difference for the 3d grade—so little as to be quite negligible. For the 4th grade there is some difference, and for each successive higher grade the difference between

the placings of the same word by the two methods becomes greater and greater as the asymmetry of the modified curves becomes more and more pronounced.

TABLE XLIV

COMPARISON OF AVERAGES OF MEDIAN DISTANCES BY NORMAL DISTRIBUTION AND BY MODIFIED DISTRIBUTIONS

	Normal Distribution		Modified Distributions	
	Unweighted Averages	Weighted Averages	Unweighted Averages	Weighted Averages
M_{3-4}	1.3326	1.3505	1.1479	1.2008
M_{4-5}	0.8471	0.8363	0.7340	0.7483
M_{5-6}	1.0406	1.0505	0.8443	0.8685
M_{6-7}	0.6344	0.6608	0.5359	0.5606
M_{7-8}	0.9201	0.9101	0.6858	0.7065

Table XLV compares the deviations from grade medians of the words of the Preferred List by Normal Distribution and by Modified Distributions. Figs. 43 to 47 give the same facts in graphic form. Words spelled by 50 per cent of pupils are of course always at 0. Words spelled by more than 50 per cent of pupils do not deviate from the median as much by modified as by normal distribution. The same is true of those spelled by less than 50 per cent of pupils. The easier a word is and the harder a word is, the greater, accordingly, is the difference in placing. The effect therefore of the modified distributions is to shorten the range of the grade scales. In using the scales, especially for pupils of the higher grades, all differences in ability between individuals or groups would tend to be decreased. It seems likely that these differences are in reality more nearly what the modified distributions show them to be. The wide range of the normal curve especially when its spread is assumed to be the same for all grades would seem to extend too far, particularly towards the low end. On the other hand, it should be said that for the words used in our scale the normal distribution gives results that are, practically speaking, satisfactory for grade scales.

TABLE XLV

COMPARISON OF DEVIATIONS FROM GRADE MEDIANS OF WORDS OF PREFERRED LIST ACCORDING TO NORMAL AND MODIFIED DISTRIBUTION. FIGURES 43 TO 47

Word No.	Word	3d Grade			4th Grade			5th Grade			6th Grade			7th Grade			8th Grade		
		% Correct	X from Grade Median / P.E. Nor. Dis.	X from Grade Median / P.E. Mod. Dis.	% Correct	X from Grade Median / P.E. Nor. Dis.	X from Grade Median / P.E. Mod. Dis.	% Correct	X from Grade Median / P.E. Nor. Dis.	X from Grade Median / P.E. Mod. Dis.	% Correct	X from Grade Median / P.E. Nor. Dis.	X from Grade Median / P.E. Mod. Dis.	% Correct	X from Grade Median / P.E. Nor. Dis.	X from Grade Median / P.E. Mod. Dis.	% Correct	X from Grade Median / P.E. Nor. Dis.	X from Grade Median / P.E. Mod. Dis.
1	even	59	— .337	— .327	79	—1.196	—1.108	89	—1.819	—1.521	93	—2.188	—1.757	93	—2.188	—1.721	97	—2.789	—2.016
2	lesson	37	+ .492	+ .482	72	— .864	— .811	83	—1.415	—1.200	91	—1.988	—1.612	94	—2.305	—1.800	96	—2.597	—1.916
3	only	65	— .571	— .552	75	—1.000	— .936	89	—1.819	—1.521	95	—2.439	—1.933	97	—2.789	—2.093	99	—3.450	—2.310
4	smoke	46	+ .149	+ .144	69	— .735	— .695	85	—1.537	—1.299	94	—2.305	—1.839	96	—2.597	—1.983	99	—3.450	—2.310
5	front	51	— .037	— .036	72	— .864	— .811	80	—1.248	—1.064	90	—1.900	—1.547	94	—2.305	—1.800	97	—2.789	—2.016
6	sure	47	+ .112	+ .108	55	— .187	— .180	69	— .735	— .634	78	—1.145	— .958	89	—1.819	—1.454	94	—2.305	—1.753
7	pear	31	+ .735	+ .721	42	+ .299	+ .286	58	— .299	— .263	72	— .864	— .728	81	—1.302	—1.048	94	—2.305	—1.753
8	bought	40	+ .376	+ .368	65	— .571	— .544	79	—1.196	—1.020	91	—1.988	—1.612	94	—2.305	—1.800	97	—2.789	—2.016
9	another	36	+ .531	+ .521	43	+ .261	+ .249	78	—1.145	— .977	86	—1.602	—1.319	94	—2.305	—1.800	96	—2.597	—1.916
10	forty	49	+ .037	+ .036	62	— .453	— .433	65	— .571	— .498	72	— .864	— .728	83	—1.415	—1.139	87	—1.670	—1.339
11	pretty	45	+ .187	+ .181	67	— .652	— .619	76	—1.047	— .894	90	—1.900	—1.547	90	—1.900	—1.515	94	—2.305	—1.753
12	wear	35	+ .571	+ .560	49	+ .037	+ .032	61	— .414	— .361	74	— .954	— .803	84	—1.475	—1.187	93	—2.188	—1.683
13	button	32	+ .693	+ .679	52	— .074	— .074	61	— .414	— .361	73	— .909	— .765	74	— .954	— .762	87	—1.670	—1.339
14	minute	26	+ .954	+ .937	38	+ .453	+ .434	62	— .453	— .394	77	—1.096	— .917	86	—1.602	—1.287	92	—2.083	—1.618
15	cousin	19	+1.302	+1.279	47	+ .112	+ .103	69	— .735	— .634	89	—1.819	—1.486	89	—1.819	—1.454	95	—2.439	—1.829
16	nails	43	+ .261	+ .319	58	— .299	— .286	71	— .820	— .705	87	—1.670	—1.373	92	—2.083	—1.648	96	—2.597	—1.916
17	janitor	19	+1.302	+1.279	42	+ .299	+ .286	58	— .299	— .263	81	—1.302	—1.083	81	—1.302	—1.048	90	—1.900	—1.497
18	saucer	11	+1.819	+1.792	29	+ .820	+ .792	42	+ .299	+ .272	58	— .299	— .257	79	—1.196	— .960	81	—1.302	—1.069
19	stopping	27	+ .909	+ .893	39	+ .414	+ .397	55	— .187	— .164	71	— .820	— .692	76	—1.047	— .838	84	—1.475	—1.198
20	sword	13	+1.670	+1.647	46	+ .149	+ .139	57	— .261	— .230	78	—1.145	— .958	86	—1.602	—1.287	93	—2.188	—1.683
21	freeze	29	+ .820	+ .806	46	+ .149	+ .139	68	— .695	— .599	83	—1.415	—1.173	86	—1.602	—1.287	94	—2.305	—1.753
22	touch	45	+ .187	+1.81	52	— .074	— .074	60	— .376	— .328	81	—1.302	—1.083	84	—1.475	—1.187	93	—2.188	—1.683
23	whistle	22	+1.145	+1.25	55	— .187	— .180	56	— .224	— .197	64	— .531	— .450	75	—1.000	— .800	85	—1.537	—1.243
24	carriage	13	+1.670	+1.647	40	+ .376	+ .360	50	000	000	67	— .652	— .550	81	—1.302	—1.048	85	—1.537	—1.243
25	nor	63	— .492	— .477	61	— .414	— .396	65	— .571	— .498	68	— .693	— .585	77	—1.096	— .877	94	—2.305	—1.753

TABLE XLV—(Continued)

COMPARISON OF DEVIATIONS FROM GRADE MEDIANS OF WORDS OF PREFERRED LIST ACCORDING TO NORMAL AND MODIFIED DISTRIBUTION. FIGURES 43 TO 47

Word No.	Word	3d Grade			4th Grade			5th Grade			6th Grade			7th Grade			8th Grade		
		% Correct	$\frac{X}{P.E.}$ from Median		% Correct	$\frac{X}{P.E.}$ from Median		% Correct	$\frac{X}{P.E.}$ from Median		% Correct	$\frac{X}{P.E.}$ from Median		% Correct	$\frac{X}{P.E.}$ from Median		% Correct	$\frac{X}{P.E.}$ from Median	
			Nor. Dis.	Mod. Dis.		Nor. Dis.	Mod. Dis.		Nor. Dis.	Mon. Dis.		Nor. Dis.	Mod. Dis.		Nor. Dis.	Mod. Dis.		Nor. Dis.	Mod. Dis.
26	already	16	+1.475	+1.453	42	+ .299	+ .286	43	+ .261	+ .238	62	— .453	— .384	44	+ .224	+ .184	77	—1.096	— .911
27	beginning	9	+1.988	+1.964	25	+1.000	+ .967	37	+ .492	+ .449	46	+ .149	+ .127	66	— .612	— .487	75	—1.000	— .837
28	chicken	49	+ .037	+ .036	70	— .778	— .733	83	—1.415	—1.200	90	—1.900	—1.547	96	—2.597	—1.983	99	—3.450	—2.310
29	choose	22	+1.145	+1.125	34	+ .612	+ .587	48	+ .074	+ .068	60	— .376	— .320	65	— .571	— .455	82	—1.357	—1.111
30	circus	20	+1.248	+1.226	39	+ .414	+ .397	50	000	000	72	— .864	— .728	75	—1.000	— .800	95	—2.439	—1.829
31	grease	11	+1.819	+1.792	18	+1.357	+1.314	37	+ .492	+ .449	35	+ .571	+ .502	42	+ .299	+ .245	57	— .261	— .242
32	pigeons	7	+2.188	+2.161	29	+ .820	+ .792	41	+ .337	+ .307	57	— .261	— .225	70	— .778	— .621	82	—1.357	—1.111
33	quarrel	15	+1.537	+1.513	39	+ .414	+ .397	53	— .112	— .099	75	—1.000	— .841	86	—1.602	—1.287	94	—2.305	—1.753
34	saucy	14	+1.602	+1.577	35	+ .571	+ .549	40	+ .376	+ .343	52	— .074	— .065	71	— .820	— .655	78	—1.145	— .948
35	tailor	38	+ .453	+ .443	55	— .187	— 1.80	70	— .778	— .669	75	—1.000	— .841	81	—1.302	—1.048	84	—1.475	—1.198
36	telegram	15	+1.537	+1.513	31	+ .735	+ .709	39	+ .414	+ .378	63	— .492	— .417	73	— .909	— .726	84	—1.475	—1.198
37	telephone	8	+2.083	+2.053	35	+ .571	+ .549	48	+ .074	+ .068	67	— .652	— .550	83	—1.415	—1.139	87	—1.670	—1.339
38	tobacco	12	+1.742	+1.716	39	+ .414	+ .397	60	— .376	— .328	75	—1.000	— .841	88	—1.742	—1.394	96	—2.597	—1.916
39	too	14	+1.602	+1.577	28	+ .864	+ .834	27	+ .909	+ .839	24	+1.047	+ .936	30	+ .778	+ .646	43	+ .261	+ .191
40	towel	24	+1.047	+1.029	44	+ .224	+ .212	64	— .531	— .461	73	— .909	— .765	78	—1.145	— .918	94	—2.305	—1.753
41	Tuesday	46	+ .149	+ .144	70	— .778	— .733	67	— .652	— .564	80	—1.248	—1.041	87	—1.670	—1.340	91	—1.988	—1.556
42	tying	44	+ .224	+ .219	58	— .299	— .286	70	— .778	— .669	68	— .693	— .585	76	—1.047	— .838	87	—1.670	—1.339
43	whole	17	+1.415	+1.392	43	+ .261	+ .249	64	— .531	— .461	78	—1.145	— .958	84	—1.475	—1.187	90	—1.900	—1.497
44	against	19	+1.302	+1.279	30	+ .778	+ .751	54	— .149	— .131	75	—1.000	— .841	84	—1.475	—1.187	94	—2.305	—1.753
45	answer	27	+ .909	+ .893	47	+ .112	+ .103	67	— .652	— .564	86	—1.602	—1.319	90	—1.900	—1.515	97	—2.789	—2.016
46	butcher	33	+ .652	+ .638	59	— .337	— .323	69	— .735	— .634	85	—1.537	—1.270	90	—1.900	—1.515	97	—2.789	—2.016
47	guess	20	+1.248	+1.226	32	+ .693	+ .668	49	+ .037	+ .034	67	— .652	— .550	77	—1.096	— .877	85	—1.537	—1.243
48	instead	32	+ .693	+ .679	48	+ .074	+ .067	62	— .453	— .394	86	—1.602	—1.319	87	—1.670	—1.340	91	—1.988	—1.556
49	raise	21	+1.196	+1.175	54	— .149	— .145	67	— .652	— .564	84	—1.475	—1.221	93	—2.188	—1.721	94	—2.305	—1.753
50	beautiful	10	+1.900	+1.875	52	— .074	— .074	70	— .778	— .669	85	—1.537	—1.270	94	—2.305	—1.800	96	—2.597	—1.916

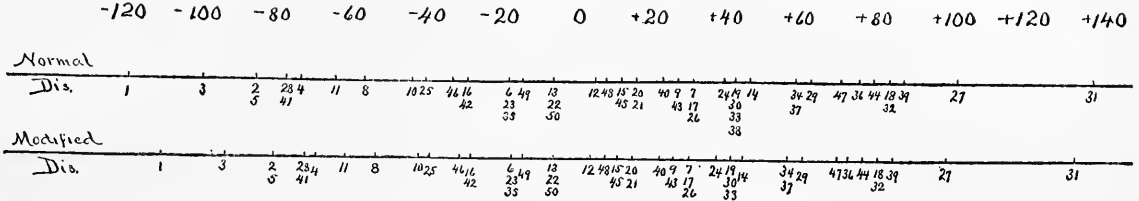


FIG. 43. 4th grade, Table XLV.

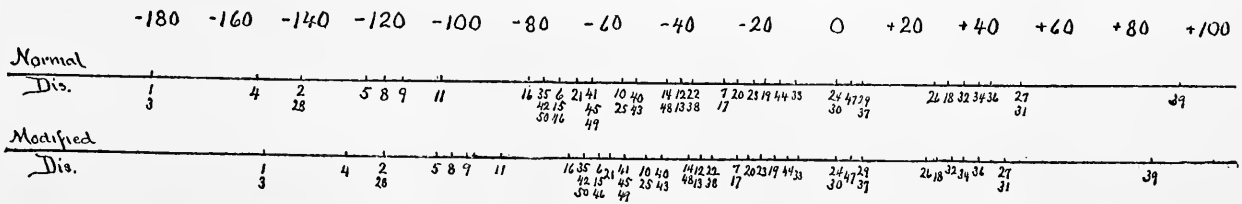


FIG. 44. 5th grade, Table XLV.

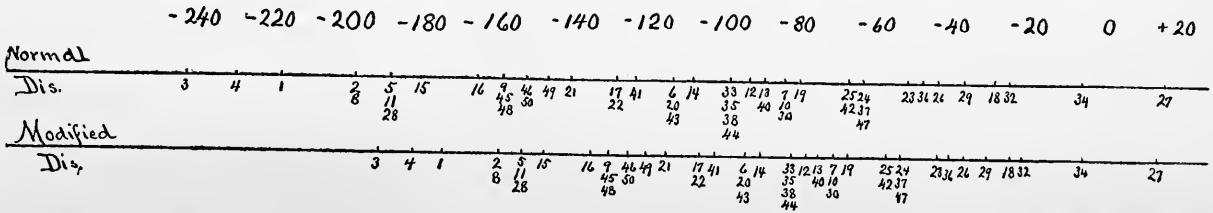


FIG. 45. 6th grade, Table XLV. Words 31 and 39 are omitted from Fig. 45. They occupy nearly the same positions by both scalings, 31 at about +55 and 39 at about +100.

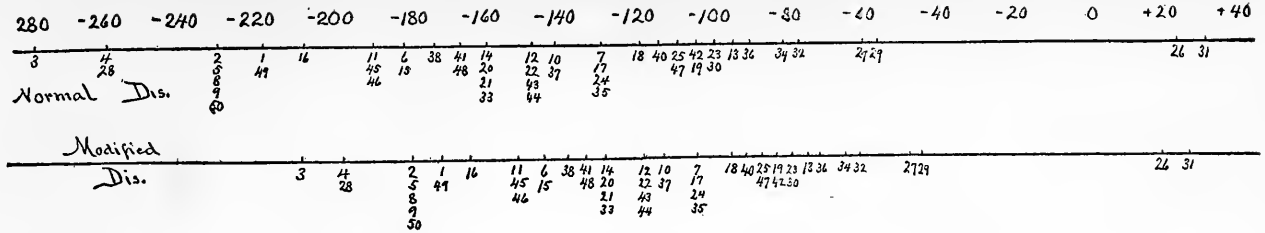


FIG. 46. 7th grade, Table XLV. Word 39 is omitted from the scale of Fig. 46. It stands at +78 on the 'Normal,' and at +65 on the 'Modified,' scale.

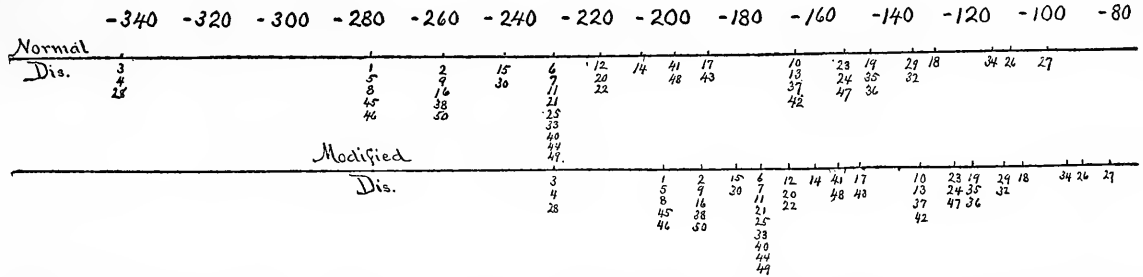


FIG. 47. 8th grade, Table XLV. Words 31 and 39 are omitted from the scale of Fig. 47. 31 stands at —26 and —24, 39 stands at +26 and +19, on the 'Normal' and 'Modified' scales respectively.

Table XLVI and Fig. 48 show a comparison for all grades combined. The same shortening of the range is evident but, whereas the contraction in the grade scales was more pronounced at the low ends, it is now in the general scale more

TABLE XLVI

THE AVERAGE POSITION OF EACH WORD ACCORDING TO NORMAL DISTRIBUTION AND ACCORDING TO MODIFIED DISTRIBUTION. POINT OF REFERENCE IS 3D GRADE MEDIAN. SEE FIG. 48

Word No.	Word	Average Position		Word No.	Word	Average Position	
		Normal Distribution	Modified Distribution			Normal Distribution	Modified Distribution
1	even.....	.699	.753	26	already.....	2.699	2.305
2	lesson.....	1.135	1.018	27	beginning...	2.917	2.525
3	only.....	.569	.604	28	chicken....	.897	.872
4	smoke.....	.835	.831	29	choose.....	2.502	2.143
5	front.....	1.057	.949	30	circus.....	2.141	1.872
6	sure.....	1.568	1.349	31	grease.....	3.294	2.838
7	pear.....	1.958	1.697	32	pigeons.....	2.739	2.378
8	bought.....	1.169	1.057	33	quarrel.....	2.069	1.816
9	another....	1.078	1.287	34	saucy.....	2.666	2.294
10	forty.....	1.758	1.477	35	tailor.....	1.866	1.579
11	pretty.....	1.311	1.137	36	telegram....	2.549	2.204
12	wear.....	1.844	1.587	37	telephone...	2.413	2.101
13	button.....	2.026	1.724	38	tobacco....	1.988	1.767
14	minute.....	1.943	1.687	39	too.....	3.491	2.998
15	cousin.....	1.681	1.491	40	towel.....	1.978	1.718
16	nails.....	1.379	1.226	41	Tuesday....	1.550	1.313
17	janitor.....	2.047	1.773	42	tying.....	1.870	1.578
18	saucer.....	2.604	2.256	43	whole.....	2.018	1.751
19	stopping...	2.213	1.894	44	against.....	2.106	1.847
20	sword.....	2.185	1.766	45	answer.....	1.594	1.425
21	freeze.....	1.740	1.517	46	butcher....	1.473	1.308
22	touch.....	1.709	1.465	47	guess.....	2.363	2.038
23	whistle.....	2.193	1.870	48	instead.....	1.756	1.517
24	carriage....	2.340	2.022	49	raise.....	1.652	1.456
25	nor.....	1.652	1.397	50	beautiful...	1.682	1.519

evident at the high end. There are also differences in arrangement, as there could not be in the grade scales. If two words which take the same position on the normal scale by ratings markedly different in upper and lower grades, but balancing each other in the aggregate, these words would not take the same position on the modified scale. The one which had

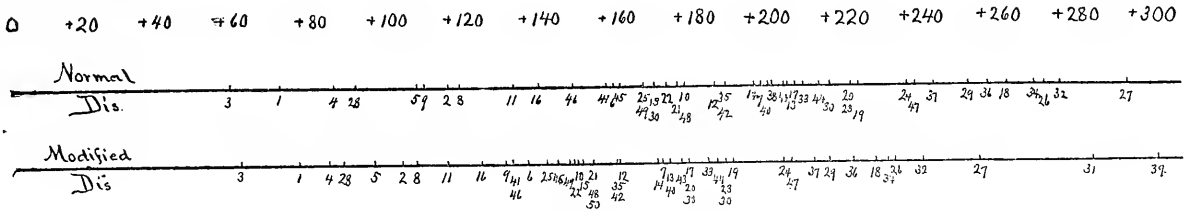


FIG. 48. General Scale, Table XLV. Words 31 and 39, at 329 and 349, are omitted from the 'Normal Dis.' scale of Fig. 48.

relatively low ratings in the lower grades would take a position above the other whose ratings were relatively high in the lower grades. This is because the modified distributions in the upper grades are such that counting in from the high end more rapidly approaches the median than does counting in the same per cent of the area of the normal curve. Take for example the words 25 *nor* and 45 *answer*. Compare the per cent ratings in Table XLV. *Nor* is easy in low grades and hard in upper grades, relative to *answer*. With the same normal distribution for all grades *nor* is a little harder than *answer*. By the modified distributions it is easier. Other words may easily be selected in Fig. 48 which show differences in arrangement. It is therefore true that the use of modified forms of distribution makes a difference which is worth noting in the scaling of words. A question to be decided on the evidence of more complete testing and a wide use of these forms of distribution is whether the differences here shown to exist impair the usefulness of the scales we have previously derived. Our judgment at present is that they do not.

§ 20. *Conclusions*

We have now certain data in hand and we may make a few general statements from them.

We have selected from a school list of about 5000 words a list for test purposes in grades 3-8 which, when put in sentences, yielded a list of 270 words. As a result of testing in two schools a selected list of 100 words was chosen and to it were added, at a later time, 18 more. These were dictated at three schools and the 100 words alone subsequently at two more schools. From the 118 were chosen two lists of 25 each. The three successive selections were made with the purpose of securing words which were easy enough in the 3d grade and hard enough in the 8th grade to afford a test in those and therefore in intermediate grades, and which showed regular increases in per cent correct from grade to grade. The two 25-word lists were then subjected to analysis and found to have high correlations between grades and between schools.

Using the entire test material and the ratings of individual pupils and assuming normal distribution and equal variability,

the differences between typical grade abilities were found and expressed as median intervals.

The 50 words which had been derived by a threefold selective process and subjected to close inspection for permanency as between grades and schools were scaled for each grade and for all grades combined. By using an "Easy 50-Word List" an expression was derived for the zero-point; and, by further testing under rigidly controlled conditions, previous grade-intervals were verified.

To fill in and extend the scale, the Rice Sentence Test was dictated and the word-scores for the Easy 50-Word List were used. It is to be understood, however, that neither of these lists was subjected to the scrutiny that was made of the Preferred List. Accordingly we cannot regard the placing of these words as very reliable.

Finally we have derived and applied tables of frequency more or less asymmetrical in character according to the amount of retention for each grade and its estimated distribution. By using them, results have been obtained which in some instances differ considerably from those obtained on the basis of a normal distribution. Such differences as appear are, we are convinced, differences in the direction of a truer representation of the facts. On the whole, however, the differences are not sufficient to impair our previous results for any practical use which is likely to be made of them.

It has become evident to us that there is a lack of knowledge of the spelling problem not only among teachers but also among those who direct their work. This is unfortunate, considering the relative definiteness of the subject and the comparative ease with which results in it may be scored. Nor is there any special consciousness of the need of more insight in this matter. Almost, if not quite, all the studies that have hitherto been made have dealt with individual performances. The behavior of words has received no attention.

It is our belief, however, that a powerful improvement in the teaching of spelling may be derived from a more critical knowledge and more accurate judgment on the part of teachers and supervisors of the material of the subject—i.e., of the words of the language. If in a list of 50 words the one word that is incontestably hardest is by more than one-fourth of a representa-

tive group of teachers judged to be the easiest, or the easiest but one, that fact in itself is a very good reason why the word is so hard. Pupils misspell it because their teachers do not realize the need of teaching it. If text-book makers disagree so widely as to put the same words in grades that are three, four, and even five years apart, it is proof of the confusion that exists as to how hard words are, and when they should be taught. There are various types of words, and each type requires different treatment. There is the type that does not need to be taught at all. There is the type which appears easy in the lower classes and (grade considered) hard in the upper classes. Such may have been prematurely taught in the lower classes. There is the type that appears to possess special difficulty for the middle grades. This is due to a constant cause—e.g., in the case of *whose,* to the learning of the use of the apostrophe in possessives. There are types of errors; there is the problem of substitution, of illegibility, and of omission.

To obtain any accurate notion of "word behavior" we must rate for words as distinct from individuals. Moreover we must give our per cent ratings thus obtained an interpretation for difficulty which takes account of the distribution of spelling ability. When we do so we shall find how unreliable percentages are as indicating differences in difficulty. We shall find, for instance, that a difference of 10 per cent between two words rated 89 and 99 means more than four times as great a difference in difficulty as is that between two words rated at 45 and 55, although the percentage difference is in both cases the same. Table XLVII (See appendix) is a ready reckoner for the conversion of percentages into units that take account of the form of distribution, assuming it to be 'Normal.'

If this study does no more than show the need of word criticism and indicate a method, it may be worth while. Every school affords a place and every day a time at which something may be done to help throw light on the nature of the material we deal with in spelling. All such work should be collected and made generally available. If teachers, principals, or superintendents who have made or who hereafter make a study of the difficulty of words, will submit them to the author of this study, the data will be gratefully received and utilized to disseminate a larger and more accurate knowledge.

APPENDIX

I. List of Authors and Titles Specifically Referred to in the Text

Thorndike, E. L. ('10). Handwriting. *Teachers College Record,* Vol. XI, No. 2.

Hillegas, Milo B. ('12). A Scale for the Measurement of Quality in English Composition by Young People. *Teachers College Record,* Vol. XIII, No. 4.

Rice, J. M. ('97). The Futility of the Spelling Grind. *Forum,* Vol. XXIII, pp. 163-172; 409-419.

Thorndike, E. L. ('13). An Introduction to the Theory of Mental and Social Measurements. Second Edition. Teachers College, New York.

Cornman, O. P. ('02). Spelling in the Elementary School. Ginn and Co., New York.

Wallin, J. E. Wallace, ('11). Spelling Efficiency in Relation to Age, Grade, and Sex, and the Question of Transfer. Warwick and York, Baltimore.

Pearson, Henry C. ('12). Experimental Studies in the Teaching of Spelling. *Teachers College Record,* Vol. XIII, No. 1.

Spearman, C. ('06). 'Foot-rule' for Measuring Correlation. *Brit. Journ. of Psych.,* Vol. II, Pt. I, July, 1906.

Brown, William, ('11). The Essentials of Mental Measurement. Putnam, New York.

Whipple, Guy Montrose, ('10). Manual of Mental and Physical Tests. Warwick and York, Baltimore.

Klein, Linus W. ('12). A Study in the Psychology of Spelling. *Journ. of Ed. Psych.,* Vol. III, No. 7.

Thorndike, E. L. ('07). The Elimination of Pupils from School. Bureau of Education, Bulletin No. 4, 1907.

Ayres, Leonard P. ('09). Laggards in our Schools. Russell Sage Foundation, New York.

Thorndike, E. L. ('10). Promotion, Retardation, and Elimination. *Psych. Clinic,* Vol. III, No. 8 and 9.

Strayer, George Drayton ('11). Age and Grade Census of Schools and Colleges. Bureau of Education, Bulletin No. 5, 1911.

Yule, G. Udney ('11). An Introduction to the Theory of Statistics. Lippincott, Philadelphia.

II. Lists Referred to in the Text and Used in the Scales

Preferred List

First

1. even
2. lesson
3. only
4. smoke
5. front
6. sure
7. pear
8. bought
9. another
10. forty
11. pretty
12. wear
13. button
14. minute
15. cousin
16. nails
17. janitor
18. saucer
19. stopping
20. sword
21. freeze
22. touch
23. whistle
24. carriage
25. nor

Second

26. already
27. beginning
28. chicken
29. choose
30. circus
31. grease
32. pigeons
33. quarrel
34. saucy
35. tailor
36. telegram
37. telephone
38. tobacco
39. too
40. towel
41. Tuesday
42. tying
43. whole
44. against
45. answer
46. butcher
47. guess
48. instead
49. raise
50. beautiful

Easy 50-Word List

1. you
2. will
3. hear
4. him
5. coming
6. he
7. is
8. on
9. the
10. road
11. and
12. almost
13. sure
14. to
15. pass
16. in
17. front
18. of
19. me
20. I
21. send
22. for
23. every
24. day
25. go
26. into
27. school
28. but
29. do
30. not
31. touch
32. table
33. also
34. has
35. only
36. one
37. pair
38. shoes
39. they
40. are
41. at
42. all
43. pretty
44. no
45. man
46. ought
47. steal
48. even
49. a
50. penny

Rice Sentence List

1. running
2. slipped
3. listened
4. queer
5. speech
6. believe
7. weather
8. changeable
9. whistling
10. frightened
11. always
12. changing
13. chain
14. loose
15. baking
16. piece
17. receive
18. laughter
19. distance
20. choose
21. strange
22. picture
23. because
24. thought
25. purpose
26. learn
27. lose
28. almanac
29. neighbor
30. writing
31. language
32. careful
33. enough
34. necessary
35. waiting
36. disappoint
37. often
38. covered
39. mixture
40. getting
41. better
42. feather
43. light
44. deceive
45. driving
46. surface
47. rough
48. smooth
49. hopping
50. certainly
51. grateful
52. elegant
53. present
54. patience
55. succeed
56. severe
57. accident
58. sometimes
59. sensible
60. business
61. answer
62. sweeping
63. properly
64. improvement
65. fatiguing
66. anxious
67. appreciate
68. assure
69. imagine
70. peculiar
71. character
72. guarantee
73. approval
74. intelligent
75. experience
76. delicious
77. realize
78. importance
79. occasion
80. exceptions
81. thoroughly
82. conscientious
83. therefore
84. ascending
85. praise
86. wholesome

III. Memorandum on the Method of Computing with Modified Frequency Tables. (Tables XXXVI-XLI.)

1. *Derivation of Median Intervals.* Table XLII, lines 4 and 5, gives for the 4th grade the number and per cent of pupils who equal or exceed the median pupil of each of the other grades. In line 6 the corresponding P.E. values are shown. These are obtained by using Table XXXVII as follows: (a) Since 80.9% of 4th-grade pupils surpass the median 3d-grade pupil, deduct 8090 cases from the high end of the 4th-grade distribution. Since there are 5236.32 above M_4(nor. dis.), 2853.68 more must be taken, extending to a point which is 1.1079 P.E. below M_4(nor. dis.). But M_4(nor. dis.) is itself .087 P.E. below M_4(mod. dis.). Correcting for this, we have 1.1949 P.E. below M_4(mod. dis.). $(-1.1079 - .087 = -1.1949.)$ This is the first entry in line 6 of Table XLII. (b) Deduct 3130 from 5236.32, leaving 2106.32. By interpellation this corresponds to +.7835 P.E. Subtracting .087 P.E. as before, we have +.6965, the second entry in line 6 of Table XLII. (c) 5236.32 less 1110 gives 4126.32, corresponding to +1.8486 P.E. Again subtracting .087 P.E., we have +1.7616 P.E., which is the third entry in line 6, Table XLII.

2. *Scaling the Words.* For "even," Table XLV, columns headed "Modified Distributions," the figures are derived as follows, using for each grade its proper frequency table: *Third Grade.* 59% correct. Count out the 5900 highest cases. There are 5141.7 above M_3(nor. dis.). We must, therefore, take 758.3 cases below that point. This brings us to —.276 P.E. Subtracting (algebraically) .051 P.E., in order to refer this to M_3(mod. dis.), we have —.327 P.E., as in Table XLV. *Fourth Grade.* 79% correct. Counting out 7900 cases from the high end, we take all the "highs" and 2663.32 of the "lows," reaching as far as —1.0212 P.E. But M_4(mod. dis.) is .087 P.E. above M_4(nor. dis.). Subtracting this amount, we have —1.108 P.E. as in Table XLV. *Fifth Grade.* When percentages are high, it is generally easier to count out their complements from the *low* end. "Even" is in this grade 89% correct. We may therefore count 8900 cases from the high end or 1100 from the low end. In either case we reach the 3245th case of the "lows," which corresponds to —1,306 P.E. Correcting for the deflection of the median from its "normal" position (.215 P.E.), we have —1,521 P.E. as given. *Sixth Grade.* $3696 - 700 = 2996$. The 2996th case corresponds to —1.339 P.E. Median displacement = .418 P.E. Subtracting from —1.339 P.E., we have —1.757 P.E., as given. The 7th and 8th grade positions are derived in the same way, care being taken to use the proper grade table of frequency in each case.

Table XLVI. The average position for each word as given in the column headed "Modified Distributions" was computed as follows: Add to the P.E. value of "even" for each grade (Table XLV) the distance which the grade median is above the 3rd-grade median. From Table XLII these distances are shown to be: M_{3-4}, 1.148 P. E.; M_{3-5}, 1.882 P.E.; M_{3-6}, 2.726 P.E.; M_{3-7}, 3.262 P.E.; M_{3-8}. 3.948 P.E. Adding these values to those of Table XLV, beginning with the 4th grade and writing the 3rd grade as given, we have the following P.E. values: —.327, +.040, +.361, +.969, +1.541, and +1.932. The average of these is +.753 P.E., as given for the word "even" in Table XLVI. The average positions of the remaining words were computed in the same way.

IV. TABLE XLVII—P.E. Values Corresponding to Given Per Cents of the Normal Surface of Frequency, Per Cents Being Taken from the Median

	0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0	.000	.004	.007	.011	.015	.019	.022	.026	.030	.033
1	.037	.041	.044	.048	.052	.056	.059	.063	.067	.071
2	.074	.078	.082	.085	.089	.093	.097	.100	.104	.108
3	.112	.115	.119	.123	.127	.130	.134	.138	.141	.145
4	.149	.153	.156	.160	.164	.168	.172	.175	.179	.183
5	.187	.190	.194	.198	.201	.205	.209	.213	.216	.220
6	.224	.228	.231	.235	.239	.243	.246	.250	.254	.258
7	.261	.265	.269	.273	.277	.280	.284	.288	.292	.296
8	.299	.303	.307	.311	.315	.318	.322	.326	.330	.334
9	.337	.341	.345	.349	.353	.357	.360	.364	.368	.372
10	.376	.380	.383	.387	.391	.395	.399	.403	.407	.410
11	.414	.418	.422	.426	.430	.434	.437	.441	.445	.449
12	.453	.457	.461	.464	.468	.472	.476	.480	.484	.489
13	.492	.496	.500	.504	.508	.512	.516	.519	.523	.527
14	.531	.535	.539	.543	.547.	551	.555	.559	.563	.567
15	.571	.575	.579	.583	.588	.592	.596	.600	.603	.608
16	.612	.616	.620	.624	.628	.632	.636	.640	.644	.648
17	.652	.656	.660	.665	.669	.673	.677	.681	.685	.689
18	.693	.698	.702	.706	.710	.714	.719	.723	.727	.731
19	.735	.740	.744	.748	.752	.756	.761	.765	.769	.773
20	.778	.782	.786	.790	.795	.799	.803	.807	.812	.816
21	.820	.825	.829	.834	.838	.842	.847	.851	.855	.860
22	.864	.869	.873	.878	.882	.886	.891	.895	.900	.904
23	.909	.913	.918	.922	.927	.931	.936	.940	.945	.949
24	.954	.958	.963	.968	.972	.977	.982	.986	.991	.996
25	1.000	1.005	1.009	1.014	1.019	1.024	1.028	1.033	1.038	1.042
26	1.047	1.052	1.057	1.062	1.067	1.071	1.076	1.081	1.086	1.091
27	1.096	1.101	1.105	1.110	1.115	1.120	1.125	1.130	1.135	1.140
28	1.145	1.150	1.155	1.160	1.165	1.170	1.176	1.181	1.186	1.191
29	1.196	1.201	1.206	1.211	1.217	1.222	1.227	1.232	1.238	1.243
30	1.248	1.253	1.259	1.264	1.269	1.275	1.279	1.286	1.291	1.296
31	1.302	1.307	1.313	1.318	1.324	1.329	1.335	1.340	1.346	1.351
32	1.357	1.363	1.368	1.374	1.380	1.386	1.391	1.397	1.403	1.409
33	1.415	1.421	1.427	1.432	1.438	1.444	1.450	1.456	1.462	1.469
34	1.475	1.481	1.487	1.493	1.499	1.506	1.512	1.518	1.524	1.531
35	1.537	1.543	1.549	1.556	1.563	1.569	1.576	1.582	1.589	1.595
36	1.602	1.609	1.616	1.622	1.629	1.636	1.643	1.649	1.656	1.663
37	1.670	1.677	1.685	1.692	1.699	1.706	1.713	1.720	1.728	1.735
38	1.742	1.749	1.757	1.765	1.772	1.780	1.788	1.795	1.803	1.811
39	1.819	1.827	1.835	1.843	1.851	1.859	1.867	1.875	1.884	1.892
40	1.900	1.909	1.918	1.926	1.935	1.944	1.953	1.962	1.971	1.979
41	1.988	1.997	2.007	2.016	2.026	2.035	2.044	2.054	2.064	2.074
42	2.083	2.093	2.103	2.114	2.124	2.134	2.145	2.155	2.166	2.177
43	2.188	2.199	2.211	2.222	2.234	2.245	2.257	2.269	2.281	2.293
44	2.305	2.318	2.331	2.344	2.357	2.370	2.384	2.397	2.411	2.425
45	2.439	2.453	2.468	2.483	2.498	2.514	2.530	2.546	2.562	2.579
46	2.597	2.614	2.631	2.648	2.667	2.686	2.706	2.726	2.746	2.767
47	2.789	2.811	2.834	2.857	2.881	2.905	2.932	2.958	2.986	3.015
48	3.044	3.077	3.111	3.146	3.182	3.219	3.258	3.300	3.346	3.395
49	3.450	3.506	3.571	3.643	3.725	3.820	3.938	4.083	4.275	4.600
50										